ROBERT HENDERSON JACK

# TANGIBILITY AND RICHNESS IN DIGITAL MUSICAL INSTRUMENT DESIGN

# TANGIBILITY AND RICHNESS IN DIGITAL MUSICAL INSTRUMENT DESIGN

ROBERT HENDERSON JACK MA(HONS), MMUS

Supervisors: Dr Andrew McPherson & Dr Tony Stockman

Examiners: Dr Thor Magnusson & Dr Chris Nash

Submitted in partial fulfilment of the requirements of the University of London Degree of Doctor of Philosophy
School of Electronic Engineering and Computer Science
Media and Arts Technology
Queen Mary University of London

Robert Henderson Jack: *Tangibility and Richness in Digital Musical Instrument Design.* Submitted in partial fulfilment of the requirements of the University of London Degree of Doctor of Philosophy © March 2019

*M*usical ideas are prisoners, more than one might believe, of musical devices.

— Pierre Schaeffer, 1977

# ABSTRACT

The sense of touch plays a fundamental role in musical performance: alongside hearing, it is the primary sensory modality used when interacting with musical instruments. Learning to play a musical instrument is one of the most developed haptic cultural practices, and within acoustic musical practice at large, the importance of touch and its close relationship to virtuosity and expression is well recognised.

With digital musical instruments (DMIs) – instruments involving a combination of sensors and a digital sound engine – touch-mediated interaction remains the foremost means of control, but the interfaces of such instruments do not yet engage with the full spectrum of sensorimotor capabilities of a performer. This poses compelling questions for digital instrument design: how does the nuance and richness of physical interaction with an instrument manifest itself in the digital domain? Which design parameters are most important for haptic experience, and how do these parameters affect musical performance? Built around three practical studies which utilise DMIs as technology probes, this thesis addresses these questions from the point of view of design, of empirical musicology, and of tangible computing.

In the first study musicians played a DMI with continuous pitch control and vibrotactile feedback in order to understand how dynamic tactile feedback can be implemented and how it influences musician experience and performance. The results suggest that certain vibrotactile feedback conditions can increase musicians' tuning accuracy, but also disrupt temporal performance.

The second study examines the influence of asynchronies between audio and haptic feedback. Two groups of musicians, amateurs and professional percussionists, were tasked with performing on a percussive DMI with variable action-sound latency. Differences between the two groups in terms of temporal accuracy and quality judgements illustrate the complex effects of asynchronous multimodal feedback.

In the third study guitar-derivative DMIs with variable levels of control richness were observed with non-musicians and guitarists. The results from this study help clarify the relationship between tangible design factors, sensorimotor expertise and instrument behaviour.

This thesis introduces a descriptive model of performer-instrument interaction, the projection model, which unites the design investigations from each study and provides a series of reflections and suggestions on the role of touch in DMI design.

# ACKNOWLEDGEMENTS

The work presented in this thesis has been supported and enriched by the help of many people.

My supervisors Andrew McPherson and Tony Stockman I cannot thank enough. Andrew has been an endless source of encouragement, inspiration and motivation throughout this PhD. Without his direction and faith in the project it never would have happened — his intellectual rigour and generosity has spurred me on throughout. Tony has been a solid support throughout this project and has given wise words of advice, always at the perfect moment. I would also like to thank Marcus Pierce for his input as internal examiner during my stage reviews. Thanks also go to my examiners Thor Magnusson and Chris Nash who through their insightful advice and guidance have strengthened this thesis no end.

It has been a great pleasure to pursue this research within the Augmented Instruments Laboratory at Queen Mary University of London, which I have seen go from strength to strength over the last few years. I would like to thank all the members of the Lab for ongoing support, being a pleasure to work with and for a four year-long stimulating conversation that I hope will go on into the future. Particular thanks go to Giulio Moro for his patience and availability that has helped this work tackle many technical hurdles, and to Jacob Harrison for being an excellent collaborator. Thank you to my colleagues on the Bela team for being as mad as required to even start on such a project – it's been a great journey in your company and long may it continue.

This research was funded as part of the wider Doctoral Training Centre for Media and Arts Technology and the minds that this programme brought together have made Queen Mary a particularly animated environment to work in. Particular thanks go to Adib Mehrabi, Olsen Wolf, Alessia Milo, and Eleanora Oreggia for accompanying me on the journey of the PhD and for always helping to keep priorities in the right order. I would like to acknowledge the anarchists mailing list which is a precious community. I want to thank everyone who has given my work their time and feedback, in particular P.A. Skantze and Nicholas Ward. Thanks to Jonathan Winfield, Geetha Bommireddy and Richard Kelly for keeping the project running smoothly, and Kok Ho Huen and Dave Wilkinson in the electronics lab.

My deepest appreciation goes to my friends, in particular my partner in crime Christian Heinrichs for keeping things emotional, and to Jonathan Rogerson, Anya Keenan, David G. Lees, Jockel Liess, Danilo Mandic, Jelena Stojkovic, Ain Bailey, Luca Nasciuti, Helen Murphy, Fred Stratford and everyone else who has been close to me while I wrote this thesis. Thanks also to my writing companions Shosh, Spock and Rosario. I thank Flora Pitrolo whose love and wisdom has guided me at every moment since we first met, and for walking this path with me. I thank the Jack pack and my wider family for their love, encouragement and curiosity throughout this project, particular thanks to my sister Rhona for providing the diagrams and to my brother David for being David. I thank my grandfather Tom Buick and grandmother Iris Jack who would have loved to see me finish. Most of all I thank my parents Valerie and Hugh for their love, trust and support for the last thirty years.

Throughout this research I was financially supported by the Engineering and Physical Sciences Research Council and Arts and Humanities Research Council via the Media and Arts Technology Doctoral Training Centre grant. Without this support I could not have undertaken this research.

# DECLARATION

I, Robert Henderson Jack, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.



*London, United Kingdom, March 2019*

Robert Henderson Jack

# PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

1. Robert H Jack, Adib Mehrabi, Tony Stockman, and Andrew McPherson. "Action-sound Latency and the Perceived Quality of Digital Musical Instruments: Comparing Professional Percussionists and Amateur Musicians." In: *Music Perception: An Interdisciplinary Journal* (2018).
2. Robert H Jack, Jacob Harrison, Fabio Morreale, and Andrew McPherson. "Democratising DMIs: the Relationship of Expertise and Control Intimacy." In: *Proceedings of the International Conference on New Interfaces for Musical Expression.* 2018.
3. Robert H Jack, Tony Stockman, and Andrew McPherson. "Rich gesture, reduced control: the influence of constrained mappings on performance technique." In: *4th International Conference on Movement Computing (MOCO'17).* 2017.
4. Robert H Jack, Tony Stockman, and Andrew McPherson. "Maintaining and Constraining Performer Touch in the Design of Digital Musical Instruments." In: *Proceedings of the TEI'17: Eleventh International Conference on Tangible, Embedded, and Embodied Interaction.* 2017.
5. Robert H. Jack, Tony Stockman, and Andrew McPherson. "Effect of latency on performer interaction and subjective quality assessment of a digital musical instrument." In: *Proceedings of the Audio Mostly.* ACM. 2016, pp. 116–123.
6. Robert H Jack, Tony Stockman, and Andrew McPherson. "Navigation of Pitch Space on a Digital Musical Instrument with Dynamic Tactile Feedback." In: *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction.* ACM. 2016, pp. 3–11.
7. Jacob Harrison, Robert H Jack, Fabio Morreale, and Andrew McPherson. "When is a Guitar not a Guitar? Cultural Form, Input Modality and Expertise." In: *Proceedings of the International Conference on New Interfaces for Musical Expression.* 2018.
8. Andrew McPherson, Robert H Jack, and Giulio Moro. "Action-Sound Latency: Are Our Tools Fast Enough?" In: *Proceedings of the International Conference on New Interfaces for Musical Expression.* 2016.

# CONTENTS

# LIST OF FIGURES

# ACRONYMS

AAF Altered Auditory Feedback
CCR Continuous Category Rating
DAF Delayed Auditory Feedback
DAW Digital Audio Workstation
DMI Digital Musical Instrument
DOF Degrees Of Freedom
DSP Digital Signal Processing
FSR Force-Sensing Resistor
GUI Graphical User Interface
HCI Human-Computer Interaction
ICMC International Computer Music Conference
IOI Inter-Onset Interval
JND Just Noticeable Difference
LMER Linear Mixed Effect Regression
NIME New Interfaces for Musical Expression
NMA Negative Mean Asynchrony
MIDI Musical Instrument Digital Interface
MSE Mean Synchronisation Error
OSC Open Sound Control
PD Pure Data
SMC Sound and Music Computing
STEIM STudio for Electro-Instrumental Music
TEI Tangible, Embedded and Embodied Interfaces
TUI Tangible User Interfaces
UDP User Datagram Protocol

# 1

# INTRODUCTION

This thesis is an exploration of the sense of touch in digital musical instrument design, and presents a series of studies analysing how touch and hearing are coupled during music-making. Learning a musical instrument is one of the most developed cultural practices based on the sense of touch, in which years of physical and theoretical training reinforce sensorimotor pathways and allow for the performance of complex music in realtime. When playing an acoustic instrument we get a great deal of rich sensory information from the parts of our body which make contact with the instrument (hands, fingertips, lips, shoulders), and it is also through those parts of the body that we physically manipulate the instrument and make it produce sound. In acoustic musical instrument performance, the importance of touch and its close relationship to virtuosity and expression is well-recognised: the sensory information received through the combination of touch and hearing contribute to temporal control [121, 275], quality judgements [93, 318, 379], expressive control [74, 196], instrument learning [3, 170, 304], and many other key aspects of musical performance. In acoustic musical practice, there is an inherent coupling between the feel of an instrument and the sound it produces.

Although touch-mediated interaction is still the primary means of control for most Digital Musical Instruments (DMIs) – instruments with digital sound generators that are separable (though not necessarily separate) from their control interfaces [221] – the instruments do not provide a rich physical experience comparable to that of acoustic instruments. With DMIs, there is a disconnect between the feel of an instrument and the sound it produces: the physical means of sound production no longer needs to be part of the control interface, and this denies the performer secondary sources of feedback about how the instrument is behaving and responding to their actions. With acoustic instruments, mappings, timbre, and the movements appropriate for control are all inscribed in the instrument's design; in DMIs however nothing comes 'for free': every characteristic of how action and sound relate has to be designed [215]. This presents a compelling problem for digital instrument design: how does the nuance and richness of physical interaction with an instrument manifest itself in the digital domain? A growing number of researchers in the field of DMI design are questioning the importance of the physical aspects of DMIs and how they relate to the digital [54, 55, 80, 91, 133, 225], and this thesis aims to contribute to this research area.

The broad aim of this research, then, is to investigate the potential *rich physical experience* of DMIs, and to understand the design parameters that influence and affect *tangible* experience (perceptual information available to the sense of touch) when performers interact with DMIs. By concentrating on the contact that a performer makes with an instrument, this research investigates an area of performer-instrument interaction that is not accessible through a consideration of hearing alone. The focus shifts to the implicit cross-modality of instrumental control, and to the manner in which action and perception are intertwined during interactions with a musical instrument.

Acoustic musical instruments have been described by Baily as 'movement transducers': the patterns and textures they create in sound are those of the movements of the body amplified [15]. In no other human artefact has the possible connection of bodily movement and sound production been as thoroughly explored: a visit to any musical instrument museum testifies to the myriad forms and evolutionary lines that were interrupted before the stable form of many instruments commonly used today was reached, and the close relationship between technological innovation and musical instrument design has been widely recognised [353]. This exploratory spirit holds into the digital context, where the latest innovations in technology and computing continue to be put to the test in the musical realm.

This thesis presents results from a series of design investigations in which new DMIs were created. Each of these instruments was designed to act as a probe that could explore a particular aspect of the sense of touch in musical interaction and monitor the effects of different design parameters on a performer's physical experience. This thesis concentrates on DMIs as a specific case of interaction with a digital system, and its main contributions are to DMI design, yet it investigates broad issues that affect Human-Computer Interaction (HCI) as much as music psychology. DMIs make interesting study cases: Tzanetakis, Fels, and Lyons have even argued that research into musical interface technology can anticipate aspects of rich multimodal and embodied interaction with computers by providing "excellent examples of sensorially rich and temporally detailed human-machine interaction" [357, p.1119].

## 1.1 PRELUDE

Before going any further I think it is important to explain how I arrived at this area of research and my initial intentions when beginning this project. In the years before embarking on this research I worked as a composer of both acoustic and electronic music and performed as an improvising musician. Having always been deeply invested in the potential of new technology for music-making (particularly in a non-realtime studio), I tried numerous times to integrate computer-based tools and instruments into a live performance scenario. Every time I tried I found I was disappointed in some way by the mediation of the computer. I would much prefer to perform with a contact microphone and series of sound-producing objects than to use the digital musical instruments and synthesisers I experimented with. This led me to question why there was such a gap between my experience with the acoustic sounding elements and those that passed through the computer: the material and physical nature of the sound produced, and the actions that produced them, were somehow disrupted by their passage to the digital realm.

The work presented in this thesis is the result of four years of investigations into how we can bridge the disconnect between action and sound in musical instruments that use a digital processor as their core; this is a broad aim that has motivated many researchers in the field of digital musical instrument research, to which I contribute my findings. This research has been conducted as part of the Augmented Instruments Laboratory where, as a group, we have developed many tools and techniques which, in various ways and with varying measures, work towards this goal. As a backdrop to this research the Lab has provided a rich source of inspiration and motivation, and the tools that we have developed together have been central to the research presented in this thesis, in both technological and methodological terms[1] and in shaping its perspective, direction and contribution.

The design of DMIs is an inherently inter-disciplinary practice that utilises techniques and knowledge from a variety of fields including musicology, HCI, design, experimental psychology and engineering. This thesis too is necessarily inter-disciplinary in nature: certain findings are of more relevance to the field of HCI, others to music psychology and perceptual studies, others to the design community. This multi-disciplinarity is something I have tried to balance as a researcher over the last four years, and something that weighs in on the presentation of my research, both in this document and in the publications that I have presented at conferences and symposia. Through crossing disciplinary boundaries I hope to illuminate issues in DMI design from different perspectives, and thus find connecting routes between fields that others may have overlooked. I believe I have achieved this in my work and that each of the contributions bring value to their fields and are strengthened through their interdisciplinary inception, both individually and as a group.

As an interdisciplinary research project this work draws from various disciplines and areas of enquiry. From perception studies I draw on work that investigates cross-modality and the connection of perception and action in musical movement. From music and musicol-

1 Most notably the Bela Platform which has been used to build all the instruments presented in this thesis and the development of which I have actively contributed to over the course of my PhD.

ogy, I draw on issues relating to performance studies and empirical musicology, organology and musical instrument research, and theories of embodied music cognition where body movement is considered as central to the perception of musical meaning. Embodiment is also the linking thread to much research from the field of human-computer interaction, in particular the position of skilled control and notions of tangible computing that question the emotive power of physical interaction with digital systems. Finally, from the field of design (a field that is already heavily inter-disciplinary) I consider the practical techniques we can use to foreground touch in the design of digital musical instruments, how these relate to notions of tangible experience and quality, and how a framework can be formed around touch in musical interaction with interactive digital systems to provide recommendations and inspiration for future designs.

## 1.2 THE FIELD OF RESEARCH

*Haptic experience*, an umbrella term for perceptions pertaining to touch [205], is becoming increasingly important to the design of all kinds of human-computer interfaces. As our everyday interactions with computers move towards more wearable, mobile and customisable devices, haptic attributes and characteristics are becoming increasingly important aspects of digital objects [100, 328]. This has led to a large increase of research into this area over the last decade. Nonetheless, the finer details of how haptic experience unfolds remains comparitively under-explored in relation to vision and audition and to the rich and well-established research areas of optics and acoustics. Technologically, also, the tools we use to engage the sense of touch are still in early stages of development in comparison to audition and vision: haptic technology remains either relatively lo-fi, generic, and with a defined function like the vibration motor on a phone, or highly specialised, expensive and research-oriented. This work contributes to the growing body of *haptics* research by focusing on the design of musical instruments, a specific case of design where touch and 'feel' are and should be of highest importance. This research also proposes that DMI designers should be treating the 'feel' and haptic characteristics of their instruments in a more methodical way and giving them more importance, instead of treating these aspects as secondary to the sonic characteristics of the instrument.

The emotional capacity of touch is central to the field of tangible computing, which calls for a rediscovery of the rich physical aesthetics of manual interaction with computers [329]. *Tangibility* itself is a rich and complex term whose definition is still a work in progress across many different fields. Tangibility is the property of an entity to be accessible to the sense of touch: physical contact of some description is central [145]. With particular relevance to DMIs, Cadoz et al.

propose two aspects of tangibility which distinguish touch from the other senses: ‘immediacy’ and ‘manipulation’ [49]. Immediacy refers to the fact that the sense of touch relies on direct physical contact, and gives us an almost intimate experience of an object that the other senses can’t provide. Manipulation, on the other hand, refers to the fact that it is through touching that we effect change on our environment and the objects within it. Musical instruments are important examples of tangible user interfaces where immediacy and manipulation take on a very specific meaning due to the high degree of skill required in musical performance and the expressive, culturally meaningful music it can produce.

Each of the studies presented in this thesis focus on the implicit cross-modality of controlling a musical instrument: the congruency and redundancy that exist between touch and hearing [337]. Perceptual attendance, that is the relative weighting of sensory information provided to different sensory channels, is crucial to the design of DMIs. Touch itself has a much lower bandwidth than vision or audition in terms of transducing information for perception [100]; however the overall amount of information available to a sensory channel does not necessarily tell us about its value or importance: the power of the haptic channel to transmit emotional (and at times life-saving) information overpowers its limited bandwidth [100]. Another way in which touch stands out from the other sensory channels is in its explicit reliance on movement, and “movement is as indispensable to touch as light is to vision” [185, p. 8].

Ideas of embodied cognition have recently had considerable influence on fields that directly concern the design of interactive audio applications such as interaction design [75], musicology [196] and digital musical instrument design [8, 80, 373]. Broadly speaking, *embodiment* concerns the presence and participation of the active body, situated in an environment, in cognition: it is through our sensorimotor system and bodily interaction with the environment that understanding arises. *Embodied music cognition* as introduced by Leman [196] proposes that music, too, is based on action: body movement is given prime importance in the formation of musical meaning, both on the side of the performer and of the audience [266]. This thesis takes theories of embodiment as its wider philosophical framing, a stance that helps clarify the dynamics of performer-instrument interaction: the close coupling of action and perception in both the performance and perception of music [120], the instrument as a mediator of musical intention [264], the manner in which an instrument is implicated by and intervenes in its surrounding infrastructure [80].

In the case of a DMI the mediation of the instrument (its translation of action to sound) does not have to be direct, and in fact much of the beauty of electronic music can be said to lie in the shift of agency over sound-making from the motor capabilities of the human to those of the machine [289]. Nevertheless, the presence of a performer's movements in the resultant sound of an instrument gives particular character to music (similarly to Barthes' notion of *the grain of the voice* [16], a way of describing the singularity of a particular singer's voice through nuances of motion). This has led Wessel and Wright [374], amongst others, to associate the tight coupling of action and sound with *expressivity* in performance. A related concept is that of *ergoticity* introduced by Cadoz [47], which posits that an essential property of instrumental interaction is the preservation of energy through both digital and physical components of a system. This is not usually the case with DMIs and it has been noted by many working within the field of computer music that the expressive possibilities of traditional musical instruments (such as the piano, the electric guitar or the cello) have not yet been matched by DMIs [54, 55, 196, 250]. Leman describes haptic feedback as "a multi-modal prerequisite for musical expressiveness" [196, p. 163] as it gives the performer a more reliable sense of how gesture translates to sound at the moment of excitation. Fundamental to this coupling is the manner in which a performer's touch is maintained in the digital system of the instrument: the congruence and redundancy between auditory and haptic perception.

Haptic engagement with a DMI is guided by static factors (material, weight, arrangement of keys, strings or frets) and dynamic factors (how it responds physically and sonically to energy put in by the performer) [250]. This research seeks to clarify the parameters of design that can help maintain the presence of a performer's touch in their interaction with a DMI, parameters that influence the feel and hence the perceived quality of a DMI and which are fundamental to how touch is catered for and understood in the design process. Perceptual attendance plays a strong role in this research: if we know the important factors of touch that performers attend to, the kind of stimuli or characteristics of stimuli that naturally stand out from the background, then we know what to simulate in high fidelity, and what we only need to give a gist of in low fidelity.

## 1.3 RESEARCH QUESTIONS

The aim of this thesis is to investigate tangible design factors in DMIs, in terms of their influence on performance and on judgements of instrument quality. Here I define the fundamental research questions, outline my approach to answering them, and clarify the scope of the work.

RQ1: How can vibratory feedback be reintroduced into a DMI and what influence does its reintroduction have on performer experience?

In particular I ask:

(a) What influence do different dynamic conditions of vibrotactile feedback have on musical performance?
(b) How accurately can musicians perform musical tasks on a DMI under each of these conditions?
(c) How does each of the vibrotactile conditions affect the musicians' impression of the instrument?

RQ2: What influence does the temporal synchronicity of audio-haptic feedback have on the perceived quality of a DMI, and how does this vary with musical expertise?

In particular I ask:

(a) At what point is action-sound latency perceptible in a DMI?
(b) What influence does action-sound latency have on the perceived quality of a DMI, in particular in relation to tangible qualities?
(c) What influence does action-sound latency have on rhythmic performance on a DMI?
(d) How do two groups of performers with different levels of expertise compare in terms of the above questions?

RQ3: How does control intimacy (the degree to which a performer's actions are reflected in the behaviour of an instrument) affect the perceived quality of a DMI, and how does this vary with musical experience?

Particular questions I address are:

(a) What influence does the level of control intimacy have on judgements of instrument quality and gestural behaviour?
(b) What influence do physical form and input modality have on performer experience?
(c) How important is the reinforcement of interaction metaphors through tangible guides in an instrument's design?
(d) How does expertise influence the above questions?

## 1.4 STATEMENT OF CONTRIBUTION

This research's main contribution is the exposition of a field of enquiry – tangibility in the design DMIs – that is currently under-explored in musical instrument design. In addition to demarcating such a field, this research puts forward a series of reflections that aim to inform current design practices and propose directions for continued work in this area. The main contributions can be summarised as follows, in the order in which they appear in this thesis:

- The vibrotactile feedback experiment in Chapter 4 demonstrates the degree to which musicians can exercise control of an instrument whilst attending to multimodal feedback.
- The effects of action-sound latency in DMIs found from the experiment in Chapter 5 are of particular relevance to understanding how small discrepancies in timing behaviour can influence the perceived quality of an instrument and the gestures of a performer.
- Findings relating to sensing richness and its role in judgements of instrument quality from the experiment presented in Chapter 6.
- Findings relating to static haptic feedback and physical form, and their influence on judgements of instrument quality from the experiment presented in Chapter 6.
- A series of reflections related to tangible aspects of DMI design outlined in Chapter 7.
- The projection model of performer-instrument interaction as outlined in Chapter 7, that can be of use in the evaluation, comparison and design of DMIs.
- A methodology that utilises technology probes in DMI research.
- A number of novel implementation strategies that involve the integration of dynamic tactile feedback into musical instruments, design for low latency, and general guidelines for building self-contained and expressive DMIs.

## 1.5 RESEARCH METHODS

This research aims to provide both technical implementation guidelines for design and theoretical contributions on the role of touch in performance with musical instruments, and therefore involves a strong practical component. The approach I have chosen involves the creation of a series of DMIs that are used as a means of testing specific theoretical territory of musical interaction.

The instruments created during this research were designed to act as 'technology probes', a concept put forward by Hutchinson et al. [150]. Technology probes are created in order to serve three goals: the social science goal of understanding the needs and desires of users in a real-world setting, the engineering goal of field-testing the technology, and the design goal of inspiring users and researchers to think about new technologies [150]. As such, each instrument was created to deliberately provoke reflection on a certain type of interaction or musical control, and to encourage thinking about the dynamics of the interaction with the instrument. These reflections were captured through video interviews and 'think-aloud' demonstration sessions. The instruments used in this thesis were also designed to gather a large set of rich information about the musicians' interaction. Each of the probe instruments deployed with the musicians were fully-tested and fully-functioning instruments that were reduced in their functionality in order to allow for a concentration on certain aspects of the interaction in isolation. The technology probe methodology is discussed in more detail in Section 3.5.

I have chosen to evaluate these instruments and their impact on the interaction using both quantitative and qualitative techniques. The quantitative methods I have employed include well-established techniques of analysis from HCI and music psychology to measure the performance of the musicians playing these instruments, evaluated along dimensions appropriate to the various particular studies presented. Some examples include measurements of instrument quality [99], measurements of accuracy of performing a certain task [273], and measurements of sensorimotor control [308]. In terms of qualitative methods, I have opted to reinforce every instance of empirical data-gathering through extensive interviews, in which performers were asked about their experience of playing the instrument and their judgements of its quality in relation to specific design parameters, and through performance observations. In the analysis of this rich data set I have opted to employ thematic analysis [40]. The particular interpretation of thematic analysis and the techniques employed are discussed in relation to its application in later chapters. In addition to these two approaches I also include my own self-reflective assessment as the instrument designer [167, 273].

All the instruments built as part of this research were created using the Bela audio and sensor processing platform. I mention Bela in this methodology discussion because the technological capabilities of the platform have enabled the creation of probe instruments in a powerful and flexible way, granting access to areas of enquiry that would have not been easily accessible otherwise. The Bela platform[2] [240] builds on the BeagleBone Black single-board computer[3] to create an embedded system that is specialised for constructing interactive audio devices. Bela is unique in its capacity to allow for hard real-time, low latency processing of audio and sensor data, in combination with high resolution analog sensor inputs (16-bit at 22.05kHz). Measurements of its performance and how it compares to other common platforms used for making interactive audio devices have been presented by McPherson, Jack, and Moro [239]. Bela also has many features that make it convenient for creating self-contained instruments: it combines the capabilities of a microcontroller for interfacing with sensors

2 https://bela.io/
3 https://beagleboard.org/black

and actuators, with the processing power of a computer and audio interface for sound generation. In each of the instrument design sections in Part ii I have provided implementation-specific details of how Bela was used.

## 1.6 structure of this thesis

The subsequent chapters of this thesis are split into three main sections; the first provides the theoretical basis of this research, surveying theories of active perception and embodied cognition, physiological details of the sense of touch and theories of sensorimotor control, approaches to tangible and tactile interactions in HCI, and the relation of this work to current research on musical instruments and their design. The second part introduces the practical element of this PhD, conducted as a series of design investigations and experiments with musicians, and contextualises this work in relation to the theoretical context outlined in the first part. The third part reflects on the findings from the practical studies presented in part 2.

Part i introduces the theoretical foundations on which this research is built and reviews the concepts of musical instrument design and haptic experience that are central to this thesis.

Chapter 2 begins by discussing the sense of touch and its dependence on movement, and moves on to focus on the special case of sensorimotor skill development when learning a musical instrument. Research that addresses the close coupling of action and perception during musical performance is reviewed, with special attention given to expression and gesture during musical performance. This chapter concludes with an overview of multimodal perception and theories of embodied cognition.

Chapter 3 brings the discussion to interactive digital systems and to how the sense of touch has been catered for in the field of human computer interaction. This begins with a review of technologies that aim to target the sense of touch, followed by a discussion on tangible and physical computing. A review of past projects that utilise the sense of touch in musical instrument design is presented. The second part of this chapter focuses on key aspects of DMI design, namely control intimacy and tangible guides and control metaphors. This chapter concludes with a discussion of evaluation techniques for DMIs and of the probe methodology used in the subsequent studies.

Part ii presents the practical elements of this research. I conducted a series of studies, each based around a new musical instrument constructed specifically for this research. Each of these investigations aims to focus on a particular aspect of haptic experience when performing with an instrument.

Chapter 4 addresses RQ1 by conducting a comparative experiment with a DMI with different programmable vibrotactile feedback patterns aimed at assisting with intonation on a single-voice instrument with continuous control of pitch. I assess the intonation of each musician while performing a series of musical tasks under each vibrotactile condition. This is combined with a gesture analysis of periods of free improvisation on the instrument with each of the conditions, and a thematically analysed structured interview.

Chapter 5 addresses RQ2 by conducting a comparative experiment with a percussive DMI with variable levels of action-sound latency. This experiment was conducted with amateur musicians and professional percussionists to assess the impact of expertise on latency perception. I assess the synchronisation error of each performer from each expertise group under each latency condition. Each of the latency conditions are directly compared to a zero latency case and rated for a series of instrument quality measures. This is combined with a thematic analysis of structured interviews conducted with each of the performers and an analysis of gestures employed.

Chapter 6 addresses RQ3 by conducting a further comparative experiment with a series of guitar-derivative DMIs that vary in levels of control intimacy (audio-driven synthesis vs. sample triggering), global physical form (guitar shaped vs. tabletop) and input modality (strings vs. touch sensor). Two groups with varying musical expertise (non-musicians and guitarists) took part. I compare the responses of each group to a pair of the guitar-deriviative DMIs with variable control intimacy, with which a series of free improvisations and musical tasks were completed. I reinforce findings from the participant responses through gesture analysis and through a thematic analysis of structured interviews.

PART III draws the first two parts of this thesis together and reflects on the findings from each of the studies.

Chapter 7 moves on to present a model of performer-instrument interaction based on projection. This model is framed by theories of embodied control and affordance structures. A post-hoc observational analysis of the gestural language employed by the professional percussionists in the study in Chapter 5 is presented.

Chapter 8 concludes this thesis by re-outlining the main contributions and revisiting the central questions and themes of this research. The main conclusions of this work are put forward and areas of future research are proposed.

# Part I

# THEORETICAL FOUNDATIONS

# 2 TOUCH, MOVEMENT AND EMBODIED COGNITION

The central questions of this thesis consider how the sense of touch affects the experience of playing digital musical instruments. This chapter serves to gather and survey previous research on human haptic capabilities and the cognitive underpinnings of sensorimotor skill, with a focus on topics that are of particular relevance to musical instruments and their design.

This thesis is motivated by a desire to understand how we can design digital musical instruments in a way that utilises the sensorimotor capabilities of the performer and provides them with a rich physical experience. In order to understand what design elements are important for this rich physical experience we have to understand how the sense of touch functions, how musical performance relates to the sense of touch, and how tangible qualities of instruments can be described.

We begin this chapter with a discussion of physiological and theoretical aspects of the sense of touch. The sense of touch is framed as inherently reliant on movement and this is put in relation to current theories of sensorimotor skill acquisition, active perception, and embodied cognition. This is followed by a more specific discussion of body movement and expressive gesture when performing with a musical instrument. The underlying work presented in this chapter lays the foundations for the practical investigations presented in Part ii of this thesis.

## 2.1 THE SENSE OF TOUCH

For an object to be tangible it must offer perceptual information to the sense of touch. There is considerable variety in the terms used to describe the sense of touch across fields, and so before proceeding any further, a clarification is necessary. What is commonly referred to as the sense of touch is in fact the product of multiple channels of sensory input: our experience of touch results from the synergetic activity of multiple distinct neural systems (often considered as submodalities of touch) that respond to various stimuli including vibration, pressure, temperature, pain, itch, joint position, muscle sense and movement [232]. *Haptic perception* is an umbrella term for perceptions pertaining to touch [205]. It can be understood as covering two distinct categories: *proprioception*, which is the sensation of the movement and position of one's body parts (closely related to kinaesthetic

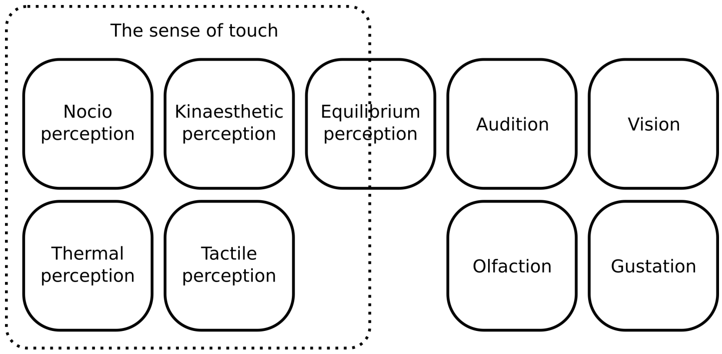



Figure 2.1: Representation of the sensory modalities brought together in the sense of touch. Adapted from Kern [177].

perception and the vestibular system that gives a sense of balance and spatial orientation for the purpose of coordinating movement), and *tactile perception* which is related to the perception of stimulation of the cutaneous receptors in the skin (vibration, pressure, texture, wetness) [100]. Additionally there is *thermal perception*, and *nociception* which is responsible for the perception of pain.

Figure 2.1 shows a representation of the different modalities brought together when we employ the sense of touch [177]. This demarcation of touch into distinct sensory channels highlights the inherent multimodality of the sense (discussed in more detail in Section 2.3) and also provides an accurate physiological distinction when considering the individual receptors at play during haptic perception [336]; however, whenever we employ our sense of touch we are in reality always using a combination of information from these sensory channels, in different ratios and weightings depending on the characteristics of the stimulus [342].

The dominant view of touch in much recent cognitive science literature has typically treated touch as inherently multisensory rather than as a single unified modality like vision and audition (see [100, 185, 205]) When employed in relation to musical interactions the complexity of the sensory modality of touch has led Cadoz to use the term 'tactilo-proprio-kinesthetic perception' [49], a slightly impractical label that nevertheless addresses the multiple sub-modalities brought together when considering the sense of touch in relation to a complex task like performing with a musical instrument. I have opted for the term 'haptic perception' to encompass this complex coupling of perceptual channels, following on from the ecological psychologist J. J. Gibson's definition of the term as "[t]he sensibility of the individual to the world adjacent to his body by use of his body" [111, p. 31]. When I choose to use the terms tactile or kinaesthetic, I am therefore referring to these particular sub-modalities of touch. Gibson's definition highlights the perceiving subject as active: it is through 'use' of the body that perception happens. Touch here is framed as based upon action: the sense of touch can be said to develop in the act of touching, that is in a human's intentional engagement with the world.

### 2.1.1 *Active touch*

*Most of the properties that we are said to perceive through touch would be imperceptible without a sense of bodily movement [...]. So touch is phenomenologically intertwined with a sense of bodily position and movement. Separating it from them would leave us with an impoverished abstraction from tactual experience, consisting of little more than base, nonlocalized sensation, perhaps with some degree of valence.*

— Matthew Ratcliffe [303, p. 137]

Perhaps the most important aspect of haptic perception is its primary role in our sensorimotor capabilities. Haptic perception encapsulates both perception and action: unlike any other sensory modality, touching is our primary means of effecting change on objects in our environment; it acts as both an input and an output action allowing us to manipulate the physical world. This aspect of the sense of touch has made it more difficult to study than other sensory modalities like vision and audition, which have historically been studied in detail in the fields of optics and acoustics; this is in part due to the fact they can be treated as sensory modalities that function through passive reception rather than active exploration. Haptics, the study of sensory information derived from the sense of touch, has also historically focused on passive reception on the skin in a similar manner to optics and acoustics [100]. Over the course of the 20th century research into haptics has increasingly investigated the active nature of touch, and the way in which action and perception are intrinsically coupled when the sense of touch is deployed [185]. David Katz, an early researcher of haptics who conducted fundamental work on the sub-modalities that contribute to the sense of touch, emphasised the role of activity in touch as well as how vision guides touching. In his book *The World of Touch* he argues that movement is as indispensable to touch as light is to vision [174]. In this respect touch can be understood as involving movement both as a perceptive act and as an action in response to perception: motor action is responsible for creating new stimuli and at the same time it is constantly modulated by sensorial information.

In his early work, Gibson too identified the importance of active exploration to the sense of touch, and proposed active 'touching' as a separate sensory modality to passive 'being touched' [111]: it is through active exploration that we are able to 'isolate invariants' in the flux of incoming sensory information [112]. Gibson demonstrates that when a subject is passively presented with a haptic stimulus (i.e. when an object is brought into contact with a stationary hand), they

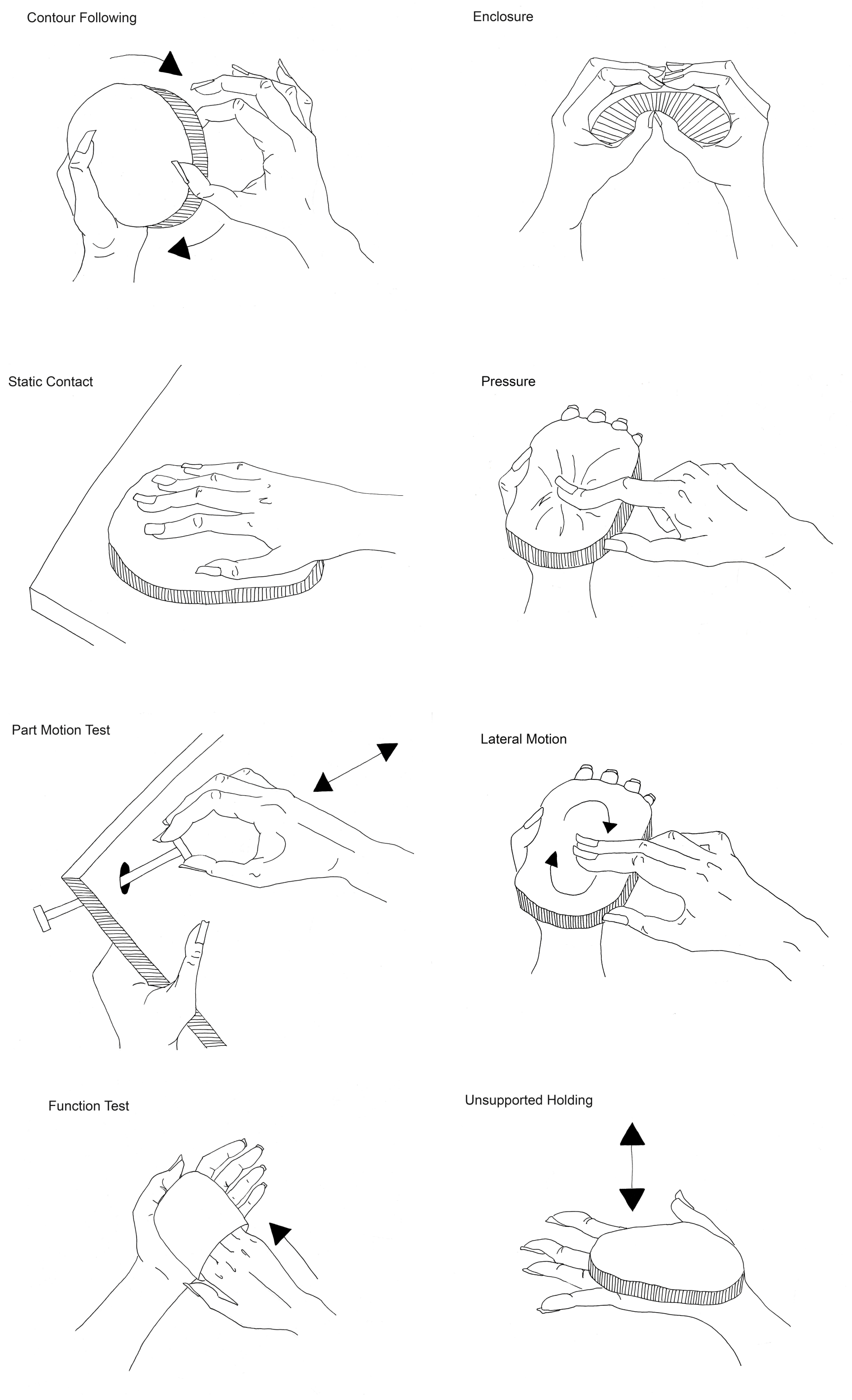



Figure 2.2: Typical movement patterns for each of the human exploratory patterns, adapted from Lederman and Klatzky [195].

will describe the object in subjective terms, for example detailing the sensations felt on the hand. In contrast, when a subject is allowed to actively explore an object they will generally report object properties and object identity [111]. Accordingly, active touch generally has higher perceptual performance due to the various human exploratory patterns tailored to each tactile property (such as texture, contour and spatial dimension recognition) [195].

As for tangible qualities – the perceptual qualities or features that are made available by the sense of touch – many dimensions of objects can be perceived haptically: for example, texture, roughness, hardness, shape and form. Lederman and Klatzky argue that these tangible qualities are bound to a certain type of exploratory hand movement (see Figure 2.2), and that there exist distinct classes of hand movements, which are directly related to distinct dimensions of desired knowledge about objects [195]. As an example they suggest considering what you would do if asked to assess the roughness of a surface. The natural response would likely be to rub the surface. Similarly if asked to assess the hardness of that same surface the natural response would be likely to use different movements like pressing into the surface, tapping upon it or squeezing it. In this manner they hypothesise that by investigating the classes of hand movement used and their relationships to knowledge goals, it is possible to better understand the representation of objects in memory and the processes by which this happens. They propose exploratory movements as 'windows' through which the whole haptic system can be viewed [195].

### 2.1.2 *A brief review of the physiology of touch*

A full discussion of the mechanisms at play in the sense of touch is beyond the scope of this thesis (see Ratcliffe [303] and Gallace and Spence [100] for reviews) and would itself be subject to considerable debate as the psychophysical and perceptual bases of tactile perception remain active fields of research. It is however worth briefly explaining the physiology of touch in terms of the mechanoreceptors involved, their characteristic behaviours, and their sensitivities.

#### 2.1.2.1 *The 4-channel model*

The sense of touch operates via a network of cutaneous receptors present in the skin. These receptors play a crucial role in our perception of pressure, texture, temperature, orientation, vibration and many other sensations that happen as a result of a stimulus making contact with our skin.

The four channel theory is one of the most established frameworks for tactile perception and explains the mechanics of four types of mechanoreceptor present in glabrous skin (non-hairy skin, i.e. lips, fingertips, palm of hands, soles of feet) [165]. These four types of mechanoreceptor differ in their depth and quantity across the body, in the type of mechanical stimuli that they respond to, and in the speed with which they respond. They are the Merkel Disk, the Meissner Corpuscle, the Ruffini Ending and the Pacinian Corpuscle [232].

A mechanical stimulation of the skin normally excites all of the receptors to varying degrees depending on the properties of the stimuli. A haptic sensation is actually a weighted mixture of four sensory channels – this makes the delineation of the individual components of haptic perception exceedingly more complex than audition, where we perceive via two copies of a single sense organ (our ears) [286]. This definition of the physiology of touch will become useful when discussing the technologies we use to stimulate this sense and the particular receptors that different techniques target.

#### 2.1.2.2 *The relationship between the sense of touch and hearing*

The potential of translating audio into tactile vibration has been explored extensively and is of particular interest to this research due to the centrality of audition and touch whilst playing an instrument. Recent research has demonstrated the great extent to which our tactile senses correspond with hearing: both are able to perceive vibrations and process frequency, waveform and amplitude [327]. Eitan and Rothschild write that both touch and audition "are based on receptors that respond to pressure stimuli, transferring them (converted into electrochemical stimuli) through the nerves to the brain for processing; and both process vibrations, analyzing (albeit with different subtlety) amplitude, frequency and waveform, within perceptual ranges and JNDs (just noticeable differences) that are often roughly compatible" [78, p. 67].

The frequency response of non-hairy skin ranges from nearly 0Hz (indentation) to around 1000Hz, in comparison with the 20Hz to 20kHz range of audition [231]. Rovan and Hayward have identified the pitch discrimination of the skin to be divided into eight to ten discrete frequency steps over the range of 80-900Hz [315]. This is just one example of a growing number of studies that evaluate how we perceive 'musical' signals through the skin. Some of the musical parameters that have undergone research in the field of psychophysics include pitch [231], timbre [316], and rhythmic acuity [183].

This research highlights the complexities of the form of sensory integration that happens between touch and hearing when playing a musical instrument: both sensory modalities, to a certain degree, perceive the same signals but with differing resolutions. It is also important to note that when it comes to the world of touch it is very difficult to generalise peoples' responses. For example it has been shown that some people like touch more than others, differ in their susceptibility to tactile illusions [134], and in their ability to concentrate on tactile stimuli [320]. This needs to be taken into account when designing custom tactile experiences.

### 2.1.3 *Sensorimotor skill*

Just as the eyes are the sensory organs of vision and the ears those of audition, the hands can be said to be the primary sensory organs of touch: although our whole body, through the cutaneous receptors under our skin and proprioceptive receptors in our muscles, contributes to the sense of touch, no single area of the human body is as sensitive to tangible qualities as the hands [286]. Alongside its sensitivity, the human hand also represents the finest example of sensorimotor skill development: a single human hand has approximately 30 degrees of freedom, thousands of receptors of varying types and sensitivities, and can provide extremely accurate and strong control without conscious guidance [303]. When we use both hands to conduct a task we also have to account for the movements of our arms and whole body: the control mechanisms involved are highly complex, yet most humans have no problem in doing so continuously as they interact with the world. These capabilities are not innate: this kind of precise control of dexterous manipulation takes years to refine.

How we learn to control these incredibly complex processes, which involve many layers of sensory processing and interaction mechanisms, is still not fully understood by the latest psychophysical research [100, 308]. Despite some areas of uncertainty there is a strong base of evidence regarding the relationship between the workings of perceptual-motor interaction (the interaction between sensory perception and active motor control of body movement) and the development of sensorimotor skill. Johansson and Flanagan [164] have examined the development of precise grasping which they identified as taking around eight years to develop. In the same paper they also suggest that sensorimotor learning consists of 3 phases: firstly, an exploratory phase, where one discovers how motor and sensory signals relate by exploring uncontrolled movement; secondly, a phase where control starts to emerge and performance improves; and thirdly, the refinement stage where performance slowly improves.

Rosenbaum, in his book *Human Motor Control* [314], highlights the importance of sequencing and timing for successful motor control and points out how every movement tends to reflect the combined use of *feedforward* and *feedback* processes. Rosenbaum points to the importance of learning by doing, learning by making mistakes then correcting, and learning by practising to the development of perceptual-motor skills. Haggard and Flanagan [127] have explained dexterous manipulation as functioning on two simultaneous temporal levels: anticipatory motor control derived from past experiences that reshape our action/perception, and contextual multi-sensorial inputs to monitor and adjust motor control.

We can summarise the action-perception loop involved in sensorimotor control as consisting of three processes that broadly relate to the feedforward and feedback modes:

- The shaping of initial action by relying on predictive models (motor memory, past experience) and feedforward perceptual cues.
- The constant monitoring of multi-modal feedback for mismatches that lead to corrective measures and sensorimotor memory updates.
- The learning of perceptual-motor skills that is mostly dependent on repetition and appropriate feedback.

These processes are employed at different points and in different measures during skill acquisition. Research on skill acquisition generally agrees that the process passes through a number of qualitatively different stages as the learner progresses from novice to expert [6, 107]. The stage-based approach is assumed to hold across different skill development domains [82].

A classic model of the stage-based approach to skill acquisition is by Fitts and Posner [87] and consists of three stages: *cognitive*, *associative* and *autonomous*. In the cognitive stage the performer has a task broken down into small components that are not related to the whole. Performance at this stage is characterised by high error, high variability and a detachment of the individual components from the whole task. The associative stage follows an extended period of deliberate practice. At this point the performer can associate actions with successful results, and thus makes fewer errors, becoming more aware of their errors through a (partial) understanding of the whole task. Once the autonomous stage is reached through further deliberate practice, the performer is able to carry out the components of the skilled action at a faster pace and without conscious attention. At this point the performer can focus on higher-level aspects of the task.

### 2.1.4 *Learning a musical instrument*

Learning an instrument can be considered a special, and highly illustrative case of sensorimotor learning where action and perception become intricately interwoven [213]. In the western classical music tradition an oft-cited and overly simplified estimate of the time it takes to master a musical instrument and become a virtuoso is given as ten years of daily practice (or 10,000 hours) [3]. Contemporary musical practices have very different understandings of what virtuosity and the mastering of an instrument mean and this amount can vary greatly from performer to performer, and also depends on the *quality* and *consistency* of the practice [333]. Nonetheless it highlights the great commitment of time and effort that it is recognised as taking to master an instrument.

In order to play music upon an instrument one needs to understand the actions that the instrument makes available and the audible consequences of these actions, an understanding that is gained through extensive practice. It is crucial to note here that the progressive understanding of an instrument by a performer is led by the performer, but grounded in the properties of the instrument itself. For this reason before discussing the development of musical skill further, it is relevant to briefly discuss musical instruments in terms of ergonomics: the study of how humans relate to the tools they use, most often applied to analyses of comfort, efficiency and safety in working environments [378].

Although musical instruments are approximately scaled to the size of the human body (an organist plays the keyboard of an organ scaled to their arms, hands and fingers, or the pedals scaled to their legs and feet, not the pipes directly; the coiled tubing of brass instruments puts the mouthpiece and valves in reachable positions; the arrangement of guitar neck and body accommodates both arms and the thighs to hold the instrument) they are often far from *ergonomic* tools when it comes to their control. Adrian Freed speaks of the *anti-ergonomy* of musical instruments[1], to highlight the manner in which they force a performer to break their body to the physical form of an instrument due to mechanical-acoustic constraints. As an example, let us consider the violin: there is little about the control of a violin from the bowing hand, to the fingering hand, to the holding of the violin, that is ergonomic or comfortable for the performer in the first instance [182]. Developing one's playing technique to virtuosic levels requires great discipline and perseverance, and for the performer to condition their body's movement patterns to be in line with the action-sound affordances[2] that the instrument offers. Rather than accommodating a performer's movements in a manner that allows them to produce meaningful results within a few minutes (as we might expect from interaction with a computer), music instruments instead require patience, practice and often physical pain.

1 https://www.youtube.com/watch?v=jbHm3DnwamM
2 This will be discussed in more detail in Section 2.4 but for now can be understood as the possibilities for deliberate action that an instrument makes available through its design.

#### 2.1.4.1 *Prediction*

As with other instances of sensorimotor skill development, extensive experience with a musical instrument leads to a strong coupling of sensory (auditory, haptic and visual) and motor processes [271] that is grounded in the physical artefact of the instrument. The realtime control of musical instruments has been noted to depend on two processes that broadly relate to *feedback* and *feedforward* models of interaction [390].

The ability to *predict* the results of an action is a key ability granted by the coupling of perception and action. If we imagine playing a note on a piano from the perspective of the musician, this involves a prediction of how their action will produce that note. As a result of the action-perception coupling in the musician's brain they are able to predict the next series of movements necessary to complete a musical task, achieved through a comparison of the current state of their body and the predicted consequences of their movements [271, 381]. This happens through a *feedforward* process where direct sensory feedback of the results of an action is not necessary for the prediction. It is only in the second stage, when the auditory and haptic feedback arrives, that the musician acts on the response of the instrument and accordingly corrects their gestures [271].

*Feedforward* control, and the prediction that this entails, are essential components of musical performance, where movements unfold in realtime and at too fast a pace for feedback to be exclusively relied upon. It has been recognised that a musician's ability to accurately and expressively perform on a musical instrument is dependent on the interaction between these feedforward and feedback models [305]. In Section 2.3 I shall present research that investigates the influence of altered feedback on musicians' ability to predict the outcomes of their movement, and hence their ability to control an instrument.

#### 2.1.4.2 *Coupling action and sound*

Knowledge of how action translates to sound is initially acquired by exploring and manipulating an instrument with somewhat arbitrary actions that lead to unexpected sounds [141]; this is what Wessel has called the 'babbling' stage due to its similarities to the manner in which young children learn to form words [373]. Through this process of experimentation, exploration and interaction, and after much repetition and the systematic association of certain actions with certain sounds, one begins to develop an *internal model* that captures the relationships of the actions the instrument affords and the resultant sound [213]. This can be understood as equivalent to the exploratory process or cognitive phase discussed above [87, 164].

An important first-hand account on the formation of this internal model comes from Sudnow [348] in his anthropological study of learning jazz piano. Sudnow describes how certain shapes of his hand came to represent harmonic structures, due the physical relationship of hand positioning and harmonic arrangements on the keyboard. Through repeated practice Sudnow describes being able to rely on the physical definition of the keys and his hands instead of having to completely plan a harmonic scale:

> I would find a particular chord, groping to put each finger into a good spot, arranging the individual fingers a bit to find a way for the hand to feel comfortable, and, having gained a hold on the chord, getting a good grasp, I'd let it go, then look back to the keyboard—only to find the visual and manual hold hadn't yet been well established. I had to take up the chord again in terms of its constitution, find the individual notes again, build it up from the scratch of its broken parts. [348, p. 12]

Sudnow describes a situation where his prediction of the movements required to achieve the next note are deeply reliant on a kind of sensorimotor familiarity with the keyboard, and on a particular set of musical theoretical structures that are built around the piano. From an initial high reliance on visual feedback, after a period of three years his playing develops to a point where he was 'going for the sounds', his musical performance became shaped by the gestural possibilities of the hand's situation on the keyboard as the coupling between action and sound gains strength:

> As I found the next sounds coming up, as I set up into a course of notes, it was not as if I had learned about the keyboard so that looking down I could tell what a regarded note would sound like. [...] I could tell what a note would sound like because it was a next sound, because my hand was so engaged with the keyboard that it was given a setting of sounding places in its own configurations and potentialities. [348, p. 45]

All available sensory information is used to build up this internal model of how actions have audible consequences with the instrument. Coupled with auditory and visual feedback, haptic feedback through an instrument's body is a natural property of acoustic instruments and essential to forming such a model. This feedback comes from static factors (material, weight, arrangement of keys, strings or frets) and from dynamic factors (how it responds to energy put in by the performer) of an instrument [250]. When playing an acoustic instrument we get a great deal of rich sensory information from the part of our body which makes contact with the instrument (hands, fingertips, lips, shoulders).

The feedback that is received as haptic energy by a performer of an acoustic instrument can be used to help them judge fine details of their gestural control, and hence influence their musical expression [279]. Leman, whose theory of embodied music cognition we will visit later in this chapter, describes haptic feedback as "a multi-modal prerequisite for musical expressiveness" [196, p. 163] as it gives the performer a more reliable sense of how gesture translates to sound in the moment of excitation.

It is also through the parts of the body that make contact with an instrument that we take control of it and manipulate its behaviour. Haptic feedback is believed to be important for the prediction and modification of sound at the millisecond level, influencing the *feed-forward* model. Haptic feedback is important to the performer as it allows a *disambiguation* of the instrument under control and its perceived effects [79] and contributes to a more reliable estimate of the sonic output of an instrument, deeply connected with musical expression and performer nuance [196].

#### 2.1.4.3 *Prediction and transparent musical instruments*

Once a certain level of expertise has been reached the performer no longer needs to focus on the playing of individual notes and can instead rely on their intuition and prediction based on the internal model that they have developed [213]. This allows musicians to think about higher-level musical constructs like phrasing, dynamics, expression rather than the coordination of their body in interaction with the instrument. Nijs [264] and other researchers who investigate embodied approaches to music cognition [196] argue that when a musician reaches this level of expertise, a symbiosis occurs between the musician and musical instrument that results from the increased integration of instrumental and interpretative movements with the musician's own body movement.

*Transparency* is the point where the performer does not need to focus attention on the individual operations of manipulating an instrument, instead focusing on higher-level musical intentions: the instrument becomes a 'natural' extension of the musician's body and so is no longer an obstacle to an embodied interaction with the music [264]. In Section 3.3.3 we shall dig deeper into what transparency may actually mean, and problematise the notion of creating natural extensions of the body, but for now we can see transparency as linked with the predictive ability of skilled musicians.

## 2.2 MUSICIANS' MOVEMENT PATTERNS

Two key questions that relate to musical communication are at the heart of music psychology. The first is about the musician and how the human body transfers an idea or mental representation into music as encoded physical energy. The second relates to the interpretation of this energy by the listener, the reverse action of transferring physical energy into action-oriented meaning, the basis of musical signification [196]. These are questions that an embodied approach to music cognition, a branch of systematic musicology, sets out to answer. The philosophical underpinnings of this approach shall be discussed in Section 2.4 however I shall now provide an overview of some areas of research that are relevant to this standpoint.

### 2.2.1 *Listening to movement*

A growing number of studies show how motor areas of the brain may be activated by observing someone else carrying out an action, by thinking about an action or even by hearing the sound of the action [116, 120]. The similarities between musical sound and body motion have been explored at length (see [120] for a review), and it is claimed that these similarities are deeply rooted in human cognitive faculties. Central to what is often termed the 'motor theory of perception' [98] is the idea that people perceive and make sense of what they hear by mentally simulating the bodily motion that is thought to be involved in the making of the sound. Godøy [120] builds upon the composer and pioneer of electroacoustic music Pierre Schaeffer's [321] concept of the *objet sonore* in order to argue that all sounds have implicitly gestural associations. This idea acts against the notion that a sonorous object is a purely abstract entity that can be perceived 'in-and-of-itself' through a process of reduced listening, a listening that leaves aside the cultural specificities or causal signification of the sound [41]. Godøy describes his belief that there is a fundamental motor-mimetic component of music perception as follows:

> There is an incessant simulation and reenactment in our minds of what we perceive and a constant formation of hypotheses as to the causes of what we perceive and the appropriate actions we should take in the face of what we perceive. I believe this points in the direction of what I would like to call a motor-mimetic element in music perception and cognition, meaning that we mentally imitate sound-producing actions when we listen attentively to music. [116, p. 318]

From this perspective, listening can be understood as involving unconscious identification of the physical origins of sound [360]. Cox's mimetic hypothesis [62] similarly proposes that when an individual hears a sound they imagine how it is created: relating to and drawing meaning from music is largely dependent on clear action-sound couplings [163]. It has been suggested that this happens unconsciously through the mirror neuron system [41]: mirror neurons are fired during both the execution *and* the observation of specific actions [310]. In the case of non-causal sounds (sounds without a clear physical origin), there is evidence that listeners associate them fairly consistently with types of gesture [52].

Leman et al. [197] conducted a study that investigated the shared basis of musical perception and musical action. They report that participants listening to Chinese guqin music display movement velocity patterns that tend to correlate with each other, and with the movement velocity patterns of the player's shoulders. These findings support the hypothesis that listeners and player share, at least to a certain degree, a sensitivity for musical expression and its associated corporeal intentionality. Peters [289], in his treaty on embodiment in electronic music, similarly suggests that musical listening cannot be divorced from the physical body of the listener, which is a conduit of every aspect of listening to sound. For Peters, even if some electronic music may tend towards a stance where the presence of the body is deliberately removed, that music is nonetheless interpreted in relation to the physical body.

### 2.2.2 *Expression and movement*

One of the key goals of musical instrument design is often stated to be the creation of tools that can foster *expressive* music-making [273]. What exactly *musical expression* means is a much-debated topic when considering the design of musical instruments [73, 126]. Fels describes musical expression as occurring when a player intentionally expresses themselves through the medium of sound [83]. The term is to a great extent bound up with the Western concert tradition: the Oxford Dictionary of Music defines it as "that part of a composer's music such as subtle nuances of dynamics which he has no full means of committing to paper and must be left to the artistic perception and insight of the executant" [176, p. 216]. Here expression is equated a performer's interpretation of a piece of scored music. It is worth returning to Barthes' notion of *the grain of the voice* [16], a way of describing the singularity of a particular singer's voice through the nuances of their vocal style and timbral variation. Both these definitions point towards the presence of a performer's movements in the resultant sound of an instrument as giving particular character to their playing.

Doğantan-Dack states that "[a]t least since the Baroque period, the *singing voice* has been regarded as the ideal model for expressive performance" [74, p. 256], introducing the notion of the 'singing hand' to describe the expectations put upon expressive piano performance. There is a danger here of conflating virtuosity and expressivity, and although the two are connected, virtuosity is not sufficient nor necessary for expressive performance [167]. The recent history of popular music brings many examples to back up this last statement; one just needs to consider punk or techno as examples of genres where quality of expression isn't necessarily the direct result of traditional musical virtuosity.

As an alternative take on what expression means, Gurevich and Treviño widen the definition in order to remove its focus on Western classical music traditions. Alongside others such as Small [334] and Clarke [59], they take a more ecosystemic approach to understanding music-making, focusing on the relationships between instruments, performers, composers and listeners as they exist within specific cultural contexts. Drawing parallels with contemporary visual arts, they discuss how expression can exist in a performer's response to the material of an instrument, rather than to a 'pre-constructed emotional complex' written by a composer that they then interpret [126]. This approach highlights the importance of the instrument, treating it as an entity that carries its own cultural context and expressive rules, rather than a tool that facilitates music. The way that an instrument physically influences the movement of performer and 'teaches them how to move' with it is another factor that weighs in on expression [370], a theme we discuss further in Chapter 7. It is important to remember that there are always a number of physical constraints at work in sound production which implicate both the instrument and the body: any musical instrument (as well as the human voice) has a limited repertoire of possible sounds it can make (though this might be very large and/or always evolving), and body motion is constrained by biomechanical and motor-control factors such as limits to speed and endurance, the need for rest or anticipation, and so forth [120].

Rebelo [304] describes how traditional African musical cultures contain a complex manifestation of expression in the making of an instrument. Natural and found materials play a significant part in this musical culture; they determine size, proportion, resonance, playability and range of an instrument. Most importantly, they assure differences between each instance of an instrument.

> [E]ach musician makes his own instrument to suit his own particular tastes. He also 'teaches' the instrument the language it will 'speak' which is, of course, the musician's own mother tongue. Bebey [19] in Rebelo [304, p. 29]

This is perhaps an extreme example of how expressivity is contained within the materials of an instrument, but it conveys the way in which expression in musical practice can be seen as relating more to participating with the instrument than mastering its control. From this perspective, expression becomes more about how the performer moves with an instrument, and equally about how that instrument conditions the movement of a performer. In Chapter 3 we shall delve deeper into a discussion of constraints from the perspective of design. Next we shall review research relating to gesture and the categorisation of human movement, discussing timescales present in musical control.

### 2.2.3 *A typology of musical gesture*

In the field of HCI there is now a large body of research that seeks to understand how human body movement can be used in the control of a computer and how the nuance and detail with which humans move can utilised. As computers are greatly reduced in their sensing capabilities in comparison to humans there is a risk that the only relevant part of a gesture becomes its 'meaning-bearing' component as judged by the computer, e.g. a key press – the rest could be considered irrelevant as it does not carry meaning-bearing information. From this standpoint the constraints of the computer can be said to define the interaction [186]: a gesture only holds value if it contains meaning-bearing content. However many identical control signals can be produced by vastly different bodily movements. Recent trends in HCI have moved to understanding gesture more broadly as an expressive stream of bodily movement, which can be recorded and analysed using motion capture technologies [163] and sensor fusion techniques [242].

Reducing musical movement to a series of meaning-bearing components does not do justice to the complex and rich sensorimotor control that musicians have over an instrument. Moving towards a more dynamic understanding of movement, with particular reference to virtual environments, Choi [58] has characterised the physiological properties of movement as a set of three *gestural primitives*: "fundamental human movements that relate the human subject to dynamic responses in an environment".

- *Trajectory-based primitives*: e.g. changes of orientation
- *Force-based primitives*: e.g. gradient movements
- *Pattern-based primitives*: e.g. quasi-periodic movements

Choi emphasises that these categories are not mutually exclusive, rather each of the three primitives can exhibit dominance in a particular movement. This purely physiological approach relates to exploratory hand movement patterns and their connection to tangible qualities as described by Lederman and Klatzky [195].

In a linguistic analysis of motion-related terminology in the proceedings of New Interfaces for Musical Expression (NIME) and related conferences Sound and Music Computing (SMC) and International Computer Music Conference (ICMC), Jensenius exposes the multitude of interpretations that the word 'gesture' can have [162]. Jensenius recommends that more care be taken to distinguish between descriptions of intention-driven action and that of physical motion. In this research efforts have been made to heed Jensenius's advice and use the terms 'motion', 'action', and 'gesture' as distinct descriptors of body movement.

The uses of the word 'gesture' in relation to music can be gathered in three main categories [163]:

- *Communication*: gestures communicate some kind of meaning in a social interaction (linguistics, psychology)
- *Control*: gestures are used in the control of computers (HCI, computer music)
- *Metaphor*: gestures project themselves on to cultural topics (musicology, psychology)

In the case of musical performance, each of these three categories are important for the interpretation of a musical performance. Gesture is usually grounded in the body of the instrument, and hence much research on musical gesture has focused on its functional behaviour, that is, on what a certain gesture achieves in a musical performance. Delalande [70], in a paper that presents an observational analysis of the playing techniques of the pianist Glenn Gould, proposes a division of the notion of gesture into three levels:

- *Effective gesture*: the movement necessary to produce a sound e.g. pressing, blowing, striking, bowing
- *Accompanist gesture*: body movements associated to effective gestures e.g. elbow movements, breathing with musical phrasing, the distance an arm is raised in relation to intensity of the note
- *Figurative gesture*: these gestures are perceived by the audience as symbolic of movement in the piece, but without corresponding to a physical movement

This thesis is primarily concerned with the movement of the instrumentalist and not that perceived by the audience; thus I will now focus more on *effective* and *accompanist* gestures than on *figurative* gestures.

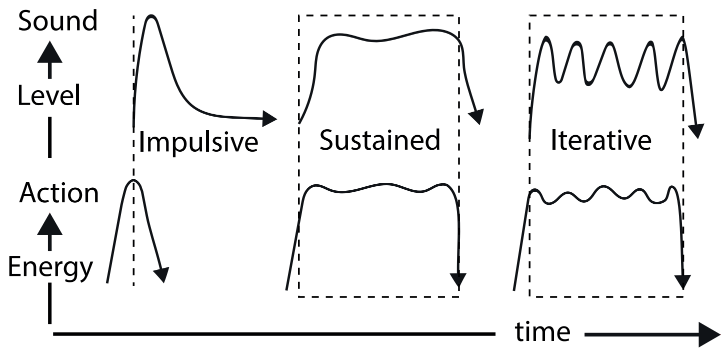



Figure 2.3: Schematic representation of the three basic dynamic typological categories of sound (above) and the corresponding motion types (below). Adapted from Godøy et al. [120].

#### 2.2.3.1 *Sound-producing gestures*

Cadoz and Wanderley [48] have further developed the idea of effective gesture as *instrumental* gesture, for which they create a typology that further subdivides it into three components: *excitation gesture*, *modification gesture* and *selection gesture*. For Cadoz and Wanderley [48] the excitation gesture is the movement that provides the energy that will eventually be perceived in the resultant output of the instrument and it can either be instantaneous (in the case of plucking, percussive strikes or key presses) or continuous (when a continuous gesture produces a continuous output or a sequence of discrete manipulations for example bowing or strumming). Modification gestures can relate to the manipulation of a sound once it is already excited, for example vibrato on a violin or pitch bend on a guitar, whereas selection gestures relate to making a choice between multiple similar elements in an instrument, such as key selection on a piano.

From a more biomechanical and motor control perspective the excitation gestures can be described using three basic dynamical typological categories of sound and their corresponding motion types (sound-motion features): *impulsive*, *sustained* and *iterative* (see Figure 2.3). The three categories of sound-motion features are distinct, yet there are thresholds where one bleeds into the other. For example when the duration of an impulsive sound is increased beyond a certain threshold it will be perceived as a single sustained sound. The important work that Godøy et al. do with this model is their connection of sound and associated motion: as can be seen in Figure 2.3 each of the action–sound types can be identified by the energy profiles of either the action or the sound, echoing the work discussed in Section 2.2.1.

#### 2.2.3.2 *Sound-accompanying gestures*

The moment of physical contact, or excitation, is only one point in a musical gesture. The parts of the gesture that bring about this contact have been called sound-facilitating gestures: for any given part of the body making contact with an instrument there a complicated kinematic system going up from the fingers, hand and arm. Ancillary gestures [367] on the other hand are gestures that are not directly sound-producing but that nevertheless are coupled to musical intention. A famous example of ancillary gesture comes from Wanderley et al. [368] in the analysis of the movement of the bell of the clarinet during performance: moving the bell isn't strictly required in order to play, but many players do it anyway as a natural by-product of playing. The intention of this thesis is not to decode what these gestures are and why they are chosen, as this has been the subject of considerable scholarship [119, 163, 351, 371]. What is important to note is that ancillary gestures themselves can serve various purposes, from being directly sound-facilitating (an extension of the sound-producing gesture), to communicative (telling something about the music, emotive, affective), to sound-accompanying (phrasing etc.).

#### 2.2.3.3 *Timescales in musical experience*

Musical experience, although it can be considered as continuous in time, depends on the perception of sound and of music-related body movement on discrete timescales, which vary from very small to very large: from the lower limits of perception (the period of a sound wave, impulses lasting a few milliseconds), to more structural features (bars, phrases, whole pieces). The various timescales at play in musical perception and performance form the basis of Godøy's [117] theory of *coarticulation* in music: the fusion of small-scale events, such as single sounds and single sound-producing actions, into larger chunks of sound and body motion, leading to what is perceived as music. Coarticulation tries to account for the distinct perceptual features that each timescale has, whether they last milliseconds, seconds, minutes, or hours, and how they are brought together as musical experience. These features can be grouped into three main temporal and spatial scale categories that relate to the sound and movement characteristics involved in each:

- *Micro*: the smallest controllable and perceivable actions, happening on a sub-millimetre scale and lasting less than half a second. For example vibrato of a finger on a violin string, trills and tremolos, and various fast transients and textural patterns, the moment of excitation.
- *Meso*: the majority of sound-producing and sound-modifying actions on musical instruments. For example moving fingers on a keyboard or changing chords on a guitar. These movements happen on a centimetre scale and typically range between half a second and five seconds. These are arguably the most important

elements for musical experience, containing the most characteristic features of style, expression and affect.

- *Macro*: large actions, moving hands, arms or full body. These movements happen on a decimetre to metre scale and last for five seconds or more. They may contain several meso-timescale objects in succession.

The *Meso* scale and dynamic characteristics of this category of features are proposed as the most important for understanding instrumental gesture. An essential concept of *coarticulation* is that this process functions on the meso-scale, but is based on continuous micro-scale elements, which concern both perception and body motion at this timescale [117]. In the next section we shall look in more detail at the temporal make-up of the micro scale and discuss some of the perceptual issues at play between the auditory and haptic sensory channels.

## 2.3 performance: a multimodal experience

Perception is an inherently *multimodal* process: all of our sensory modalities mutually influence each other as we derive meaning from the environment [342]. When we perceive we are always dealing with information from various sensory modalities in different weightings and ratios. As an example we can imagine hearing a car horn coming from behind. The most likely response to this would be to turn and see where the sound is coming from. In this reaction at least three sensory modalities are employed: the auditory, visual and vestibular. We are alerted to something happening behind by hearing the sound of the horn, we then move to direct our limited line of sight in the direction of the sound, in moving we employ the vestibular system to guide body movements by informing us of the position of limbs and head and to keep balance. In everyday life we are always deriving meaning from our surrounding environment with all the sensory information we have available [337].

### 2.3.1 *Audio-haptic feedback in musical instruments*

It is worth reiterating here that a distinction can be made between two types of touch: *active* and *passive*. Gibson proposed that active 'touching' be considered a separate sensory modality to passive 'being touched' [111]. Active touch brings together information from tactile and kinaesthetic receptors during the active manipulation of an object. Passive touch, on the other hand, occurs when stimuli are presented to a stationary person [100]. Wollman et al. [380] note that in the case of the violin both active touch and passive touch are inherently involved: active touch is mainly involved in the bowing of the strings and the movement of the fingers whilst fingering on the strings, while passive touch receives vibrotactile feedback from the resonating body of the instrument through the parts of the body that make contact with it. Differentiating between these two modes of touching is useful in terms of specifying the particular type of touch we are describing: in practice however, the two are coupled in everyday perception.

Haptic feedback through an instrument's body is a natural property of most acoustic instruments due to the intrinsic link between control mechanism and sound production. Askenfelt and Jansson [12] have shown that the vibrations of four acoustic instruments (double bass, violin, guitar and piano) can readily be perceived by their performers for almost all positions of the instrument in normal playing. They suggest that this feedback plays a role in the timing of musical gestures, and in ensemble playing (where hearing one's own auditory feedback is not always possible), concluding by stating that perhaps the major feature of an instrument's vibration is "to convey a feeling of a resonating and responding object in the player's hands" [12, p. 347].

The connection of haptic cues and the reported quality of an instrument has long been known by acoustic instrument makers: in the case of the piano, a key's material and its dynamic response from its coupling to the hammer mechanism give a piano a unique haptic signature [175]. In a series of experiments conducted in a Leningrad piano factory in the 1970s Galembo and Askenfelt showed that blindfolded and 'deafened' pianists were able to easily identify three pianos they had previously played when presented with one of the three at random. Interestingly, the group of pianists performed worse at piano identification when they were only able to listen to them being played [99]. This study illustrates the close connection between our understanding of instrument identity and the unique feel of an instrument, suggesting that the role of touch is more important to the recognition of a specific instrument than hearing.

To further investigate this question Wollman, Fritz, and Poitevineau [379] conducted a study to assess the influence of vibrotactile feedback on quality assessments of violins by violinists using an isolated violin neck with a vibration transducer attached. They found increased vibration resulted in enhanced ratings for the following criteria: 'rich sound', 'loud and powerful' and 'pleasure'. In a follow-up study they compared the role of both auditory and tactile feedback on similar quality measures [380]. They found an influence of both auditory and vibrotactile feedback on the violinists' evaluations but reported fundamental differences in the way violinists interpreted and evaluated each of the quality measures. Similar perceptual studies, again in relation to the piano, have been conducted by Fontana et al., again finding increased preference across quality measures for conditions with additional vibrotactile feedback [90, 91] but with results tempered by the difference between individual performers and their preferences.

Chafe suggests that tactile feedback is most valuable for a performer by helping them to determine when a note is settled and stable on an instrument [55]. This echoes Askenfelt and Jansson's [12] proposal about self-monitoring in ensemble performance. Fulford, Ginsborg, and Goldbart [96], in a review of the influence of hearing impairments on musical performance, report that double-bass players can check their tuning via haptic beat frequencies between their pitch and that of the ensemble by means of the haptic 'beating'[3] felt in the instrument's body. Relying on haptic feedback for intonation may be particular to bass instruments due to the frequency range of the instrument (well within the threshold for tactile perception over most of its range), the comparatively large area of neck over which beating may occur and the slow frequencies at which it happens, and the nature of the musical material, which tends toward long pedal tones where there is time to receive this kind of feedback.

In terms of temporal control, Goebl and Palmer [121] show how tactile feedback is used by some pianists to increase timing accuracy in piano playing, particularly at higher tempi. This growing number of studies points towards the critical role that haptic feedback plays in instrumental control. In a follow-up study with clarinettists the importance of kinematic landmarks for temporal accuracy was observed [281], leading to the suggestion that the amount of sensory information available at finger-key contact can enhance the temporal accuracy of music performance.

In a study conducted by Finney [86] in which pianists played a series of pieces on a keyboard with and without auditory feedback, it was found that there was no negative influence of missing sound on the accuracy of their performance. Finney found no significant differences between the normal and silent conditions with regard to number of errors, overall tempo, overall dynamic level (average key-press velocity), between-hand asynchrony, and variability of note durations [86]. This suggests that in the case of the piano haptic feedback alone is enough for pianists to monitor their performance [86]. A similar study identified expressive components of pianistic performance – horizontal and vertical timing, horizontal and vertical dynamics, and pedalling – to not be dependent on auditory feedback [305]. Interestingly the same does not seem to be the case for string players, where audio masking has a negative impact on intonation. In one of the few studies that investigates audio feedback deprivation on an instrument that is not the piano, Chen et al. [57] have shown that cellists' intonation gets significantly worse when performing shifts of position in the absence of auditory feedback. This is likely due to the continuous,

3 A more in-depth description of beating can be found in Section 4.4.

rather than discrete, pitch scale that makes string players more reliant on richer feedback. This topic is one of the points of investigation in the study presented in Chapter 4, which investigates how vibrations in an instrument's body can guide the intonation of a performer.

### 2.3.2 *Audio-haptic integration*

The study of interactions between stimuli concurrently presented to different sensory modalities is now a central area of research in psychophysics due to its importance in spatial perception [342]. It has been posited that it is in the congruency and redundancy between information presented to different sensory modalities that ecological meaning lies [79, 337], i.e. in deriving meaning from all available sensory information. It has also been widely recognised that multisensory conditions of stimuli presentation offer a more ecologically valid situation compared to the uni-sensory conditions of stimulation that characterise many previously published psychophysical studies, for example where passive tactile perception of stimuli has been the focus [308]. With more sensory channels involved, experimental design necessarily becomes more complex. In what follows we shall review research that particularly targets the temporal make-up of audio-haptic perception. This is of particular relevance to the study presented in Chapter 5 where I investigate the influence of action-sound asynchronies on impressions of an instrument.

#### 2.3.2.1 *The window of simultaneity perception*

*Multisensory integration* is the process by which the human nervous system merges available sensory information into unique perceptual events [50]. Temporal correlation between stimuli is the basis by which our brain integrates information from different sensory channels into a single event, and also differentiates between sensory information that is and is not related to an event. An important measure is the point of subjective simultaneity: "the amount of time by which one stimulus has to precede or follow the other in order for the two stimuli to be perceived as simultaneous" [339, p. 365]. This threshold varies quite substantially depending on the individual and on the sensory channels involved. The distance between the stimulus area and the brain also has an effect on this threshold. Joining stimuli received through separate sensory channels can take place between stimuli that are temporally asynchronous, but which fall within the "temporal window" of integration [245] which for audio-haptic stimuli can vary from tens to hundreds of milliseconds depending on various factors to do with the location, magnitude and content of the stimuli [275].

Tapping studies are useful for examining how movements are synchronised with an auditory stimulus and help enrich our understanding of sensory pathways and the weighting of signals in cross-modal perception. Levitin et al. [202] and Adelstein et al. [2] investigated the perceptual asynchrony threshold values for an active audio-haptic interaction situation (playing a drum). When the auditory stimulus either preceded or followed the strike by between ±200ms a threshold value of -25ms and 42ms was reported by Levitin, whereas for Adelstein et al. thresholds varying between 18ms to 31ms depending on the stimulus duration were reported in an experiment with audio only being delayed. Both these studies, amongst others [307], report that some participants had very low threshold values (ca. 10ms), particularly musicians. Musicians have been identified as having lower audio-tactile simultaneity perception thresholds than the general population due to their high level of auditory-tactile training [2].

### 2.3.3 *Simultaneity perception*

Delayed feedback (be it auditory, visual or tactile) can cause disruption to the nuanced sensorimotor control used with instruments. In the field of HCI delayed feedback has mostly been studied as system latency (the asynchrony between a control gesture and a system's corresponding response) and jitter (the variability of this asynchrony). Latency and jitter are fundamental issues affecting interactive digital systems and have long been recognised as potentially harmful to a user's experience of control [210, 243]: even if accuracy of performance is not affected, the qualitative experience of the user may be negatively impacted [171]. This is discussed in greater detail in Chapter 5.

#### 2.3.3.1 *Altered auditory feedback and musical performance*

The Altered Auditory Feedback (AAF) paradigm, in which the sound that results from an action is altered [293], has been used extensively in music psychology to investigate the importance of auditory information for the execution of control sequences on musical instruments. Delayed Auditory Feedback (DAF) is a common form of AAF where the onset of auditory feedback is delayed by a fixed amount in relation to the action that produced it [32]. While the feedback is usually kept as what a performer would expect in DAF, with only the synchrony of perception and action being affected, there are other types of AAF that alter the contents of auditory feedback while maintaining synchrony. For example, experiments have been conducted on digital keyboards where the AAF consists of shifting pitches to disrupt expectations of pitch arrangements on the keyboard [291, 294, 296] or randomising pitch [86, 292].

Each kind of alteration to auditory feedback disrupts performance in different ways and to different extents. Recent research on delayed feedback suggests that asynchronies between action and feedback primarily disrupt the timing of actions, not their sequencing (the production of melodic sequences) [291]. Pitch alterations on the other hand disrupt the accuracy of production and not timing [291]. The point of maximal disruption caused by asynchronies (the amount of delay, above which no significant increase in disruption is seen) has been the focus of much research. Generally disruption increases as the delay increases up to a certain point, and then reaches asymptote (Gates, Bradshaw, and Nettleton [103] found an asymptote around 270 ms in music performance). However, rather than an absolute time discrepancy, the degree of disruption caused by asynchronies depends on when it occurs in the Inter-Onset Interval (IOI) in rhythmic performance and reflects the phase relationships between onsets of auditory feedback relative to the IOI between actions (key presses for example) [293]. For example in case of tapping along to a metronome beat at 120bpm (IOI of 500ms), if a delay of 250ms is introduced to the auditory feedback, this will affect performance less than an IOI of 300ms which does not have a regular phase relationship with the IOI of the target beat. This disruption varies with delay length and similar effects have been shown for DAF in speech [147].

#### 2.3.3.2 *The sensorimotor conflict hypothesis*

A common interpretation of disruption from delayed feedback is the sensorimotor conflict hypothesis. The proposal is that delayed feedback interferes with the planned timing of actions [293] or their execution [147] due to shared representations for perception and action [209]. Delayed feedback causes disruption by conflicting temporally with the expected timing of a planned movement [295], in this case the expected sound that should result from an action. The magnitude of disruption most likely depends on the perceptual salience of the delayed feedback [343].

Our concern is with audio-tactile interactions: the tight coupling between auditory and tactile feedback systems has been recognised [275], as has its increased temporal resolution of synchrony perception in comparison to audio-visual and tactile-visual [95]. Whereas the importance of auditory feedback for musical performance is evident, given the primary aural focus of music as a cultural practice, tactile feedback has been shown to play an important role in the control of timing during music performance [122], and expert performers have been shown to depend less on auditory feedback and more on tactile feedback than non-expert performers during the performance of sequential movements [340]. High temporal acuity is shared by both hearing and touch. In terms of temporal precision hearing is the most accurate of our senses: two stimuli of equal subjective intensity were perceived as being temporally discrete if separated by ca. 2ms for monaural and binaural stimulation [202], touch being less accurate (ca. 10-12ms [110]) but still better than sight (ca. 25ms [178]).

#### 2.3.3.3 *Musicians' timing ability*

Musicians have been recognised as better than non-musicians across a range of timing dependent tasks. In duration-based tasks where the identification of the duration of two intervals are compared musicians outperform non-musicians [302]. Musicians also show a superior ability in distinguishing timing changes within isochronous sequences [203], which is particularly true of percussionists who demonstrate the highest accuracy of all musician groups [77]. The average error when tapping to an isochronous sequence is also notably smaller for amateur musicians in comparison to non-musicians (10-30ms vs 20-80ms) [9, 307].

Further differences between instrument speciality have been demonstrated: participants with high levels of rhythm-based musical expertise (in particular percussionists) display superior synchronisation abilities (smaller average error and less variability in tapping tasks) when compared to other musicians and non-musicians [51, 184, 222]. Dahl [65] reported that professional percussionists demonstrated a variation of mean synchronisation error of between 10-40ms which equated to 2-8% of the associated tempo. Even lower synchronisation error in professional drummers has been reported by Kilchenmann and Senn [179] (between 3ms and 35ms depending on motor effector, part of the drum kit and rhythmic 'feel') and Hellmer and Madison [137] (below 5ms).

### 2.3.4 *Perceptual attendance*

Temporal simultaneity is one aspect of sensory integration, another being the weighting with which we give stimuli received simultaneously through different sensory channels, and the effects that they can have on one another (known as congruency and redundancy effects [342]). When trying to explain which modality dominates over another and under what circumstances this occurs, the modality appropriateness hypotheses by Welsh and Warren are often cited [34]. The basis of these hypotheses is that discrepancies in stimuli are always resolved in favour of the modality that provides the most precise representation. For example, in the case of a spatial task, such as reaching for a glass on a table, the visual modality usually dominates, because it is the most precise sensory modality at gathering spatial information. When temporal judgements are required audition tends to dominate, when textural perception is required, the haptic sensory channels are the most precise and thus dominate.

In order to better understand the mechanisms of multi-sensory integration, psychology experiments have long focused on the places where sensory modalities come in to conflict or create illusions. The McGurk effect is perhaps one of the most famous examples of intermodal conflict, where what is being seen changes what is being heard [233]. In terms of auditory-haptic integration there have been a number of studies conducted on the influence that audition can have on what is felt on the skin. Hötting and Röder [146] conducted a series of experiments where a single tactile stimulus was delivered to the right index finger of subjects, accompanied by one to four task-irrelevant tones. They found that all of their participants reported feeling significantly more tactile stimuli when two tones were presented simultaneously than when no or only one tone was presented. Bresciani et al. [42], in a similar experiment, investigated whether the perception of tactile sequences of two to four taps delivered to the index fingertip could be modulated by simultaneously presented sequences of auditory beeps when the number of beeps differs (less or more) from the number of taps. Their results showed that tactile tap perception can be systematically modulated by task-irrelevant auditory inputs.

Again these studies focus on the passive perception of tactile stimuli, rather than active tactile interaction. Guest et al. [123] conducted a study that focused on active interaction and surface texture perception where participants judged the roughness of abrasive surfaces which they briefly touched. The sounds of the participant touching the surface were captured by a microphone and presented back to the participants through headphones in three different conditions: no processing; amplified; attenuated. The participants rated the surfaces on two different scales: smooth-rough and moist-dry. They found that attenuating high frequencies led to an increased perception of tactile smoothness (or moistness), and conversely the boosted sounds led to an increased perception of tactile roughness (or dryness). This has also been explored by Tajadura-Jiménez et al. [350] in relation to touch screen interaction.

#### 2.3.4.1 *Perceptual hierarchy*

Gallace and Spence [100] provide an insightful anecdote that highlights the hierarchy according to which we attend to perceptual information. The example is of an immersive Formula One driving simulation featuring wrap-around high definition screens, a real Formula One driver seat, and near-realistic controls. They speak of how upon first opening, the public were generally disappointed by the promised 'immersive' aspect of the simulation. The addition of a fan with the rotation speed of the blades mapped to the speed of the car greatly improved the public's response to the simulation [139]. This anecdote serves to show how wise use of low-fi sensory feedback in combination with other stimuli can greatly improve the whole sensory experience.

Spence and Gallace propose that the complete and accurate simulation of individual complex perceptual stimuli is often not the best route to creating a sense of immersion in interface design – capturing the user's attention with a simple stimulus is often more effective. As we only pay attention to a small part of our environment at one time [337], it is not necessary to accurately simulate that which falls outside of the focus of our attention. In other words, by knowing what people are going to attend to, or what kind of stimuli naturally stand out from the background, it is possible to know what to simulate in high fidelity, leaving the rest in the background in low fidelity as a gist of the stimuli. This line of argument is important when considering the sensory integration of touch and audition when performing with a musical instrument.

The mechanism of sensory integration is complex and depends on many factors; yet as these studies show, sensory integration is at times very stable in its behaviour, and certain effects can be easily recreated between participants. This has led to the exploitation of such effects to create what has been termed pseudo-haptic feedback (see [194] for a review). Stable multisensory effects can be used to modulate the perceived haptic qualities of an object [311]. In the studies conducted as part of this research, presented in Part ii of this thesis, I am interested in finding similar stable territory: effects of sensory integration of haptic and auditory stimuli that have a distinct and defined behaviour within musical interaction. We shall now move on to discuss embodied cognition, a set of theories that have become of increasing importance to our understanding of interaction with artefacts and tools that extend our ability to express ourselves.

## 2.4 EMBODIED COGNITION

Over the last few decades there has been a shift in cognitive science, computing, robotics, psychology and philosophy towards theories of 'embodied cognition' that treat perception and cognition as deeply dependent on the physical body of an acting human [330]. This shift towards embodiment is, amongst other things, a reaction to Cartesian dualism, whose mind-body split dominated philosophy and cognitive psychology from the 1600s onwards. The Cartesian perspective simplifies psychological processes by equating them with symbolic computations within the brain: cognitive science hence limited its investigations to what happens inside the head [330]. Cognition from this point of view begins with an input to the brain and ends with an output from the brain. This model was reinforced by developments in computer science in the 1950s-60s and equates cognition to an 'information system', where there is a strictly unidirectional flow from sensory perception (input), to cognition (processing), to action (output) [230].

As an alternative, theories of embodied cognition developed as a group of loose-knit theoretical stances in cognitive science and phenomenological philosophy that share a commitment to critiquing and moving forward from traditional approaches to cognition and cognitive processing [330]. The general unifying principle of embodied theories of cognition hinges upon an understanding that perception and action are closely intertwined and can mutually influence one another: the human body, with its perceptual and action capabilities, and in interaction with the outside world, became the central focus of theories of human cognition. The roots of embodied cognition can be traced back to the phenomenology of Husserl, Heidegger and Merleau-Ponty, and to Wiener's theory of cybernetics [377] which explored the structuring of intentionful action and the role that feedback plays in non-symbolic cognition. Varela, Rosch, and Thompson's reinterpretation of the phenomenologists' work as 'enactive cognition' in the book *The Embodied Mind* was particularly influential. In this book Varela, Thompson and Rosch describe enactive cognition and its relationship to embodiment as follows:

> By the term embodied we mean to highlight two points: first, that cognition depends upon the kinds of experience that come from having a body with various sensorimotor capacities, and second, that these individual sensorimotor capacities are themselves embedded in a more encompassing biological, psychological, and cultural context. [...] [T]he enactive approach consists of two points: (1) perception consists in perceptually guided action and (2) cognitive structures emerge from the recurrent sensorimotor patterns that enable action to be perceptually guided. [361, p. 172]

Rather than treating body movement as a mere outcome of higher-level symbolic and rule-based manipulations, the enactive view emphasizes the role of *sensorimotor* engagement in cognition: perception and action are deeply intertwined.

### 2.4.1 *Perception as a form of action*

For Merleau-Ponty [246] perception requires action: all senses are active and directed at gaining knowledge about our environment. This echoes the material discussed in Section 2.1.1 and has been extended as the sensorimotor contingency theory [265]. This theory posits that conscious perception happens as a result of movement of the sense organs in combination with a low-level understanding of how the sensory input received should change based on the movement of that sensory organ. This theory also argues that the sensory modality that best helps to understand perception in general is touch, not vision, as has traditionally been the case (for example [224]). The close coupling of perception and action represented by the sense of touch, as both capable of sensing and acting, has served as a paradigmatic example for many theorists who now think of perception as inherently active. Noë [265] for example, defends an account of perception understood as essentially a form of action. He begins his account by arguing that vision is 'touch-like', involving sampling and exploratory probing of the environment. Perception in these terms, instead of being a passive act of receiving sensory information, becomes a 'skillful bodily activity' [265] in itself.

### 2.4.2 *Body, environment, affordance*

Although much of the groundwork for embodied cognition came from phenomenologist schools of philosophy, similarly influential themes have been explored in psychology. The aforementioned psychologist J.J. Gibson, who is of particular importance to this research due to his acute insight into the mechanisms of the sense of touch, spent much of his later career concerned with visual perception [111], gradually becoming disenchanted with common understandings of visual processing in psychology which separated perception from action. Gibson's main argument is that perception is not purely information processing, but hinges on the relationship between living creatures and the environment within which they act. Although this theory was developed in relation to visual perception, it is applicable to perception more generally and was developed as a result of Gibson's earlier work that places heightened importance on movement in the functioning of the sense of touch [111].

An essential aspect of the theory of ecological perception is that perception, and the particular information picked up through perception, is specific to the *action-capabilities* of the perceiving subject: perception is directly predisposed to the actions of an organism in the world. Gibson initially termed this predisposition as environmental *affordance*: "what [the environment] offers the animal, what it provides or furnishes, either for good or ill" [113, p. 127]. Affordances denote the possibilities for action that an object holds in a specific situation which exists relative to the action capabilities of an actor. For example a guitar affords certain action possibilites to a guitarist and a different set to young child. Equally a set of stairs affords a different set of action possibilities to a mobile adult than to a toddler [372].

Others such as Norman [267] and Gaver [105] have since found ecological psychology, and particularly the notion of affordances, a valuable tool for analysing and designing interactive systems. This approach has been widely accepted in the fields of HCI and design, so much so that Norman's concept of *perceived affordances* – the properties of an object that define the actions a user perceives as possible – is now widely understood and used by designers [267]. Design here is understood as providing affordances to the user of the designed object. Theories of ecological perceptions have also been applied to listening [106, 341], and extended to theories of musical listening by Clarke [59]. Clarke suggests that understanding meaning formation in musical listening involves taking our everyday listening and the capacities of our auditory system as the point of departure, echoing Gaver's work on the perception of everyday sounds [106]. We shall return to this discussion in Section 3.5.5, when we will look in more detail at DMI design.

### 2.4.3 *Implications of embodiment*

Ideas of embodied cognition and enaction have had significant influence in fields related to music and to the design of interactive systems. Some of the most notable contributions are to interaction design [76, 181], musicology [3, 196], music pedagogy [170], and digital musical instrument design [8, 80, 373] as shall be discussed further in Section 3.4.

#### 2.4.3.1 *Embodied music cognition*

In Section 2.2.1 we introduced a series of experiments that illustrated the close relationship between music perception and body movement which generally fall under the umbrella of embodied music cognition. Embodiment when applied to music cognition means that listening to music becomes not just a matter of processing sound input, rather it becomes about the reenactment and mental simulation of the body motion associated with the sounds that are heard. One premise of this approach is the idea that music is an inherently multimodal phenomenon: music is not treated as an exclusively aural sensory process, but rather as one that pertains to all the senses [196]. Furthermore, in order for the cognition of music to be embodied it must be based on the capacities and limitations of the body in relation to the environment. An embodied approach to music cognition therefore suggests that there is an inherent coupling between action and perception: corporeal imitation and movement are essential aspects of musical expressiveness.

Considering music from the perspective of embodied cognition has motivated a great deal of research in musicology and related performance studies [196, 213, 264, 266]. The majority of studies that happen under the banner of embodied music cognition involve empirical data sets derived from the likes of performance tracking using motion capture systems: measurements of movement becomes the focal point of this form of musicology, as covered in Section 2.2.1. There is also an increasing number of works in musicology that are less directly measurement-based, and which provide subjective accounts of a performer learning or mastering a musical instrument from an embodied perspective. Notable examples of this are Le Guin's account of embodiment in the cello works of Boccherini [193], and Aho's account of learning the Kantele, a Finnish folk instrument [3]. Both of these works echo Sudnow's account of learning jazz piano, referenced earlier in this chapter [348].

#### 2.4.3.2 *Embodiment in HCI*

In HCI terms, an embodied perspective does not necessarily translate into a set of generic guidelines for design or a framework. It should rather be understood as an alternative perspective on human-technology interaction [349] where the focus is shifted to direct manipulation, the coupling of technology and human action and the way that technologies relate to the sensorimotor capabilities of the body. Dourish [75] argues that in order to manage the meaning of an interaction we have to better manage this coupling.

In terms of interaction design this can translate to more deep consideration of the interaction techniques that allow for rapid coupling between user action and the reaction of a system. Svanæs states that "[i]n order to allow for fluid integration into the perceptual apparatus of the user, the action-reaction coupling should be one that is easily 'understood' by the body" [349, p. 8:26]. With this shift comes a renewed importance given to the finer details of how technology influences movement, of physicality and skilful control, and of thinking through doing [181].

Svanæs [349] reminds us that although all interaction with technology is at base embodied, the relevance and importance of the embodied perspective increases as the *proximity* between technology and the body increases. Proximity here is not exclusively meant in a physical sense, rather it is about the tightness of the coupling between technology and the body. The more closely integrated the lived body of the user and the technology become, the more relevant it becomes to consider this interaction from embodied perspectives. In Section 3.3 we shall discuss control intimacy, a central concept of this thesis that relates to this coupling. Additionally, in Section 3.2 we discuss the field of tangible and physical computing, a field whose goal is a rediscovery of the embodied basis of HCI.

THE 'FEEL' DIMENSION The embodied perspective places us in a good position to consider the haptic and tangible aspects of interaction with technology, which Larssen, Robertson, and Edwards [192] term the 'feel' dimension. They describe this dimension as the process of using the kinaesthetic and haptic senses when incorporating a tool into the bodily space. Characterising the feel dimension in interactions with technology requires a consideration of the *experiential* aspects of a user's control of a device, an aspect explored at length by Tuuri, Parviainen, and Pirhonen [355] who, in a recent paper on embodied control within HCI, discuss human-technology interaction. The focus of their discussion is on how movement is used in the control of technology, and reciprocally, how movements are controlled by technology, thus constituting what they term 'technology-induced choreographies'.

In this paper they highlight how everyday actions are continuously choreographed by both our natural and technological surroundings. These choreographies can be understood as consisting of action-affordances; perceived possibilities for action defined by the environment [113]. They also highlight how the cultural significance of an artefact can shape its choreography: "[t]echnological artifacts and user interfaces do not exist in a vacuum. Rather, they are inevitably fused into the context of a more general physical and social infrastructure that facilitates, triggers, guides and orientates the dynamics of everyday movements" [355, p. 2]. The theories put forward by Tuuri et al. shall be revisited and expanded upon in Section 3.6 when we introduce a model of performer-instrument interaction.

## 2.5 CHAPTER SUMMARY

This chapter has discussed theories of cognition and human sensory capabilities that relate to touch and musical performance. It began with a discussion of the sense of touch and its reliance on movement, focusing on the hands, the perception of tangible characteristics, and the acquisition of sensorimotor skill.

Learning a musical instrument is a very specific case of sensorimotor skill acquisition, and Section 2.1.4 focused on the processes at play when we gain expertise with a musical instrument. Through repeated practice a coupling between action and perception is formed which allows performers to act in *feedforward* mode, predicting the outcomes of their actions. This enables realtime performance, and in combination with the monitoring of feedback from the instrument, forms the basis of musical performance.

Section 2.2 looked in more detail at the types of movements that are employed in musical performance, how they have been categorised and analysed, and how musical expression can be understood in relation to movement. This section also reviewed studies that point to the close relationship of musical experience and body movement: listening to music becomes not just a matter of processing sound input, rather it becomes about the reenactment and mental simulation of the body motion associated with the sounds that are heard.

The micro timescale of musical interaction was the focus of Section 2.3. Through a review of research from music psychology and psychophysics that focuses on the temporal make-up of audio-haptic perception. This research, as well as giving an example case of the close coupling between the two sensory modalities, serves as background for the study presented in Chapter 5.

The final section discusses the larger philosophical framework of embodiment which places human movement at the centre of cognition. Theories of embodied control in human-computer interaction are central to the discussion of interaction with DMIs that is presented in Chapter 7.

# 3

# TANGIBLE CONTROL AND DESIGN

*Every instrument has its difficult and easy fingerings, its rough and smooth terrain. A singer's effort in reaching a particular note is precisely what gives that note its beauty and expressiveness. The effort that it takes and the risk of missing that note forms the metaphor for something that is both indescribable and the essence of music.*

— Sally Jane Norman, Michel Waisvisz and Joel Ryan [270, n.p.]

We now move on to discuss technology and design in relation to the human haptic capabilities that have been explored in Chapter 2. This chapter focuses on the design of interactive digital systems, with particular attention given to the sense of touch as it relates to DMI design. It begins with a discussion of HCI and its relationship to sensorimotor skill. This is followed by a review of technologies specifically designed to provide haptic experiences in interactive digital systems, and a discussion of recent trends in the much broader field of tangible and physical computing, where the affective potential of interaction with tangible elements of a digital system has been explored.

We then return to the specific case of the musical instrument, and review current research and design investigations that relate to the following key subjects: incorporating haptic feedback; control intimacy; tangible supports for musical interaction. This chapter concludes with a discussion of evaluation techniques for DMIs and introduces the projection model of performer-instrument interaction. By contrasting approaches from design, HCI and music psychology I outline the challenges and opportunities of evaluating such devices.

## 3.1 DESIGNING FOR SENSORIMOTOR SKILL

The history of HCI can be understood as driven by ongoing attempts to utilise the full range of human skill and ability in the control of computers [75]: by exploiting the rich sensorimotor control patterns that humans naturally display with other objects in their environment, *natural mappings* can be created between an intention-driven movement and the corresponding reaction of a computer. We shall revisit the notion of 'natural' mappings in a later section, but the first section of this chapter focuses on the physical devices we use to input information into a computer, and on how movement is translated into an approximate representation, as a trajectory or series of discrete events that can be understood by the computer.

The field of HCI can be described as progressing in three waves since first gaining popularity in the early 1970s with research into visual computing and Graphical User Interfaces (GUIs); speaking of these three waves helps identify generational trends in HCI research. Each wave builds upon the foundations of the previous ones and does not nullify the findings and approaches from those before [131], rather they can be thought of as global shifts in perspective within the field.

The first wave is generally characterised as an exploration of human-machine coupling, heavily inspired by industrial engineering and human-factors and ergonomics. This wave aimed to improve the 'fit' between humans and machines. In terms of nomenclature, whereas the first wave focused on *humans* and *machines*, the second wave (1990s) switched the focus to *users* and *computers*: models of information processing systems became the dominant paradigm for understanding interaction with computers, driven by the wide-scale introduction of personal computers in the workplace and at home and by the kinds of interaction that this technology instilled [356]. The information processing model can be best described as consisting of three parts: the input signals that are captured through the senses; the processor that transforms or alters this information; the output of the system, whether that be a thought, action or response. Central concerns of this wave were how information is taken in, what transformations it undergoes, how it comes out, and how this can be done as efficiently as possible.

The nascent field of HCI has since grown to encompass a broader view of interaction that can better address research questions that did not easily fit into the information processor model, such as how a particular interaction feels to a user, how an interaction with a computer fits into and has influence on everyday life and more ethnographic questions about the wider social implications of a technology [131][1]. The third wave of HCI (2000s) introduced *embodied interaction* as an underlying theme. Embodiment, in terms of considering the human body as central to understanding interaction, was implicit to the two prior waves, whether in the form of assessing the fit of a mouse to the hands of a user, or the readablity of a particular font size: human-factors and ergonomics. Equally studies on the limits of human sensorimotor ability, for example the speed with which a particular task could be completed, or reaction times to changing stimuli, were central to second-wave HCI [319].

Embodiment, as reincarnated in third-wave HCI, draws more on phenomenology and on the manner in which our understandings of

---

1 The technology of an era also has great influence on the way we understand cognition. Metaphors of the mind have always moved with the latest technological advances, from the book, to the filing cabinet, to the computer file structure, to the distributed network of the internet [39].

the world, of ourselves and of our interactions with the world derive crucially from our location in a physical and social world and our active engagement with it (echoing the material discussed in Section 2.4). A focus on embodied interaction shifts the goals of HCI and of what we consider as central to interaction [131]. Klemmer, Hartmann, and Takayama [181] review theories of embodiment from sociology, psychology and philosophy in order to offer a series of themes that they see as particularly relevant to interaction design. They identify a shift from the second wave's idea that thinking is cognitive, abstract and information-based to one where thinking is active, achieved through doing things and expressed through gestures, and where cultural factors, meaning-making, and the influence of technology on everyday life take on increased importance [181].

### 3.1.1 *Input devices – from the physical to the digital*

Intention-driven body movement is understood by a computer as a control signal via the physical manipulation of an input device. In most cases this depends on dexterous manual interaction: to control computers we most commonly use our hands, and although there is a move towards voice control of computers this is still limited in terms of realtime applications. When reviewing modes of manually inputting information into a computer Bill Buxton [46] humorously comments on the rudimentary and reductive nature of the input devices in use in 1987. He imagines a scenario where a future physical anthropologist was trying to understand human physiology and physical capabilities based on the equipment they found in a computer shop of the time. He states that they might think that we had a "well developed eye, a long right arm, uniform length fingers and a low-fi ear. But the dominating characteristic would be the prevalence of our visual system over our poorly developed manual dexterity" [46, p. 366]. Although the variety of commercial input devices on offer has increased over the last 30 years, Buxton's parable still partially rings true.

The input devices that have underpinned the history of HCI have almost all involved manipulation led by the hands: buttons, dials and sliders, the keyboard[2], the mouse, multi-touch track pads, motion tracking of the hands. These input devices register human movement in different ways, with what Verplank [362] describes as buttons and handles (discrete and continuous sensing) being by far the most common sensor typologies. With each technique there is a discretisation or dimensionality reduction of movement into a signal that a computer can understand: actions become bounded, thresholded and

2 Early versions of the typewriter were based on the piano keyboard and the two have shared a long interlinked history [199].

quantised onto set trajectories. This can result in information being discarded from the interaction.

For the most part of human history our relationship to tools has been direct: the human operating a tool had to produce the required force and control needed to operate it. In this case the operator of the tool received feedback through their body directly and immediately. Slowly, through windmills and water turbines, and then through the great advances in mechanisation that the industrial revolution brought about, the power sources and control mechanisms of different kinds of tools could be disconnected from one another, abstracting control gestures from the operation of the machine [136]. The direct feedback loop that existed was broken, giving rise to the control interface.

Input devices do not only serve as means of transmitting information to the computer (although this is of course their primary goal) they also contain other information about the quality of an interaction in the way they behave, informing a user's impression of control. In their review of the 80-year history of industrial design from the company Danfoss, which mainly produces controllers for central heating, Øritsland and Buur [276] reflect on some of the aspects of user experience that they see as having been lost during the upgrade from electromechanical devices, to analog electronic devices, and finally to digital devices:

> We are concerned that interaction designers in enthusiasm with new technologies fail to preserve or transfer the qualities of use which were achieved with outdated technologies. For instance, the digital adjustment of settings using plus/minus buttons, though more precise, lose the feeling of being-in-control and the sense of range and proportion offered by analog potentiometer knobs. [276, p. 27]

Øritsland and Buur identify the gradual reduction in the richness of interaction and physical feedback given to the user. With the digital device, movement has been reduced and simplified to actions on a very small scale: finger presses on buttons and feedback from an LCD screen. With 'enhanced' control which increased the functionality being made available to the user (in this case the ability to program central heating on an hour-by-hour basis) came a loss of appreciation for sensorimotor skill: with the digital device very different functions can be controlled with very similar actions [72]. This shift in the type of control we have over technology can be seen across a wide range of consumer products and devices, and is also true of a whole generation of digital music technology which is highly reliant on generic control interfaces mapped to synth engines, nested menu structures, and complex control sequences of simple button presses to access features. Indeed, the explosion in popularity of hardware modular synthesisers that are mostly based on analogue electronic technologies could in part be seen as a reaction against the digitalisation and intangibility of software-based musical instruments [283]. This discussion touches upon issues of the importance of physical experience in our relation with devices and products, a topic that we shall discuss further in Section 3.2. For now it is worth identifying that Øritsland and Buur touch upon a fundamental aspect of HCI which regards the separation of control and operation.

### 3.1.2 *Technologies of touch*

There has been continual growth in research on the sense of touch and its use in HCI from the 1960s to now. The idea of purely haptic interfaces was born from this research, and a generalised model of a haptic interface can be seen in Figure 3.1 and can be described as follows:

> Unlike traditional interfaces that provide visual and auditory information, haptic interfaces generate mechanical signals that stimulate human kinesthetic and touch channels. Haptic interfaces also provide humans with the means to act on their environment. We can therefore attempt to define haptic interfaces as being concerned with the association of gesture to touch and kinesthesia to provide for communication between humans and machines. [135, p. 16]

A driving force of early technological developments in haptics was the reintroduction of haptic feedback into situations where it had been lost through mechanisation which separated control from operation. An example of this is the control stick of a plane, which had become decoupled from the control mechanism of the plane through developments in electronic aeronautical technology: this was potentially dangerous as, in the absence of direct mechanical feedback, pilots felt disconnected from the movement of the plane and the forces it was under [100]. Similar areas of research centred on communication through the skin; for example in the form of instructions provided to individuals working in hazardous conditions, or in conditions with limited visual information (e.g. fire fighting or during military operations) [211]. Another dominant trend in early haptics research was *sensory substitution*, which explored how the signals received through one sensory channel could be translated to another; an example of this was blind people being provided with visual information through stimulation of the skin [309]. Despite large research efforts and economical resources, the practical applications of haptics research at large in commercial products remain few and far between in comparison to developments that target vision or audition [101]. The most common type of haptic feedback to be widely implemented

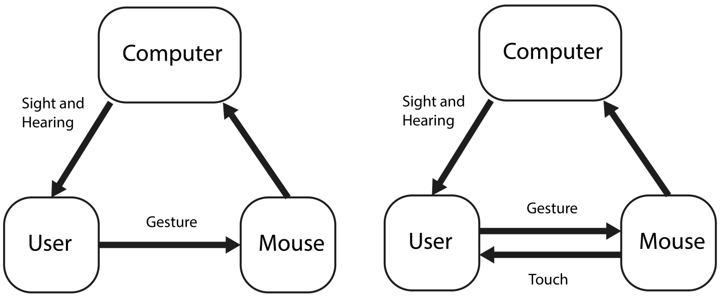



Figure 3.1: Model of a haptic interface adapted from Hayward et al. [135].

in a product is the vibrate function on mobile phones. In terms of complexity in relation to the capacity of the sense of touch, the vibrate function on most phones is equivalent to a flashing light for vision or a single tone alarm for hearing.

Tactile feedback itself – that is sensations received through the cutaneous receptors under the skin – has been shown to act as a support and source of confirmation when interacting with multi-touch interfaces [391] and both dynamic and static tactile feedback have been shown to aid navigation and orientation on touch screen devices when a user is able to actively touch them [297]. Some of the experimental devices that have emerged from this research field include the tactile television, which enabled users to detect basic images that were stimulated on their backs [60], stockmarket figures represented through a vibrating belt that allowed for a perception of 60 words per minute [138], the promise of a new sense that provided the position of magnetic north through a vibrating belt to assist users in navigation [257].

Applications for the deaf and hard of hearing have also been developed, exploring how sound could be represented through tactile vibrations on the skin. The tactile vocoder was developed as an aid for lip-reading, and represented the spectrum of speech as an array of vibrating stimuli equally spaced across the forearm [43]. Early implementations investigated the translation of sound to passive tactile perception of periodic vibrational stimuli on the skin. In its simplest form, a tactile stimulus that takes advantage of the skin's sensitivity to vibrations can be created by sending an audio signal to a small loudspeaker placed on the fingertips which acts as an actuator. This form of direct haptic monitoring of an audio signal is similar to 'speaker listening' as widely practised by deaf and hard of hearing people over the last century [132]. Using full-range audio signals to drive actuation is not the ideal solution, as there are important differences between the tactile sensory channel and the auditory sensory channel as discussed in Section 2.1.2.2.

### 3.1.3 *Haptic technologies for DMIs*

The lack of haptic feedback in electronic musical instruments has been acknowledged by many as a central problem in their design [36, 55, 227, 272, 315]. With the dissociation of control mechanism and sound production come limitless possibilities for translating action into sound and relating a performer's movement to the behaviour of an instrument's sound. However, a strong and rich relation with the concrete body of the instrument is lost. In recent years much attention has been given to the possibility of reintegrating this rich kinaesthetic and tactile feedback into instrument design through the use of electronic actuators. Vibrotactile feedback, the production of vibrational stimuli usually achieved through mechanical actuation, aims to recreate the vibratory behaviour of an instrument as it resonates. There are three main categories of commonly used vibrotactile actuators which can be seen in Figure 3.2. Each is suitable for different applications in tactile interface design:

ROTARY ELECTROMAGNETIC ACTUATORS (VIBRATION MOTORS) Rotary direct current motors are perhaps the most common vibrotactile actuators used in consumer electronics. They are designed to rotate a shaft continuously when a constant current is applied to the motors. An off-centre mass is fixed to the output shaft of the motor so that its rotation exerts large radial forces on the body of the motor. Frequency and amplitude are coupled to the motor's rotational speed. There is also a delay time in start up due to the spinning of the shaft and attached mass (the delay time is in the region of 200ms). They are most commonly used for notifications, for example in mobile phones and pagers, and are also found in the 'rumble packs' of many game controllers [165].

(a) Vibration motors (b) Voice-coil (c) Piezoelectric

Figure 3.2: Three commonly used actuator types.

LINEAR ELECTROMAGNETIC ACTUATORS (VOICE COILS) These are essentially electromagnets where a varying current creates a varying magnetic field and the core component of both loudspeakers and microphones. In a voice coil actuator a permanent magnet will either be attracted or repelled from the magnetic field. This principle is used in speaker drivers to produce the movement of the speaker cone. This type of actuator has a rapid response time and can output a large range of frequencies depending on size and design [165].

NON-ELECTROMAGNETIC ACTUATORS (PIEZOELECTRIC) Vibrotactile sensations can also be created by taking advantage of the piezoelectric effect, wherein particular solid materials change shape when subjected to an electrical current. These actuators also respond very quickly and can output arbitrary waveforms, however they have a limited amplitude and often require high voltages [165].

In recent years there have been some promising developments in in-air haptic feedback that uses arrays of ultrasound speakers to stimulate the skin [53]. Due to the experimental nature and cost of this technology, it falls outside the scope of this research.

#### 3.1.3.1 *Integrating vibrotactile feedback*

Actuators should be chosen according to the desired frequency response, dynamic response and their planned arrangement. The mounting material and method is of prime importance to the effectiveness of the actuator and to whether the vibrotactile stimuli it produces are diffused or direct. General principles for integrating haptic feedback into a DMI have been outlined at various points in the field. In their guidelines for audio-driven vibrotactile feedback Rovan and Hayward [315] suggest developing different typologies of tactile sound events, which relate to the sound of the instrument in terms of dynamic temporal behaviour and spatial positioning. The division they propose is as follows:

- **Type A – Time dependent correspondence**: Use to signify discrete events, either caused by the performer or machine. Example: switching program states, zone borders.
- **Type B – Space dependent correspondence**: Use to guide continuous movement. Example: relation to absolute position, speed of entering a zone, sharpness of a "whip" or hitting motion, virtual friction of "bowing" movements.

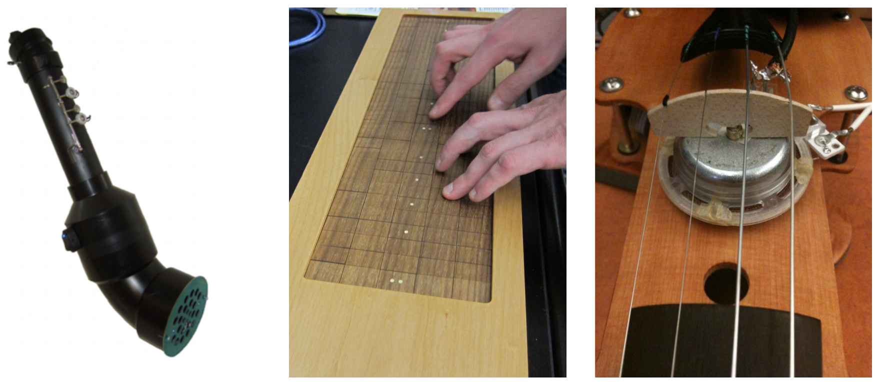
Birnbaum and Wanderley [30] suggest that tactile feedback should always be targeted at specific cutaneous receptors and their perceptual frequency range and sensitivity to particular types of vibration, through actuator choice and Digital Signal Processing (DSP) techniques. Giordano and Wanderley [114] have explored the extent to which vibrotactile feedback can be driven with signals that contain behaviours that are equatable to musical audible signals. Marshall and Wanderley [228] have explored how tactile feedback can influence performer engagement and perceived control of a DMI.

VIBROTACTILE FEEDBACK IN DMIS Vibrotactile feedback has been added to many DMIs (see Figure 3.3). Overholt [278] augmented a traditional violin to create a violin-family instrument that incorporates sensing, embedded DSP, and physical actuation of the acoustic body via an actuator that was also responsible for producing the acoustic sound of the instrument. This was with the aim of creating an implicit link between the tactile and acoustic behaviour of the instrument. Similarly, Marshall and Wanderley have explored the addition of instrument-like vibrotactile feedback on a performer's engagement with a DMI [228]: the Viblotar is a DMI which can couple its sound output to structure-born vibrations that can be felt in the performer's knees and hands. They found increased participant enjoyment for settings with vibration feedback, but a negative effect of the feedback on participants' perceived ability to control the instrument.

Birnbaum and Wanderley created the BreakFlute and Touch Flute, two flute-like instruments with added vibrotactile feedback provided by actuators mounted in the tone holes that targeted the fingertips of the performer [30]. The acoustic signal of the flute is analysed in realtime to derive pitch, onset and loudness measures that are then used to drive an envelope generator which produces vibrotactile pulses of a particular pitch and duration. Birnbaum and Wanderley here are trying to stay close to instrument-like vibrotactile feedback but are tuning specific parameters of the feedback (frequency range, timbral pallet, intensity) to make it more in-line with the capabilities of the tactile modality. This builds on the work of Rovan and Hayward who proposed mapping strategies and DSP techniques for vibrotactile feedback in relation to an in-air instrument. One of the most ambitious projects to reintroduce tactile feedback comes from Papetti, Schiesser, and Fröhlich [282] who created an expressive musical interface with multi-point vibrotactile feedback by adding piezo disc based actuators to the surface of an H-Plane produced by Madrona Labs.

#### 3.1.3.2 *Haptic Force-feedback*

An alternative take on introducing haptics to DMI design has targeted the kinaesthetic and proprioceptive aspects of touch: those that relate to awareness of one's body state, including position, velocity and forces of the body in motion. Haptic force-feedback controllers are designed to target exactly this modality and provide resistance to the movement of the part of the body that is attached to them (usually a finger or a hand).



(a) BreakFlute [30] (b) H-Plane [282] (c) Overtone fiddle [278]

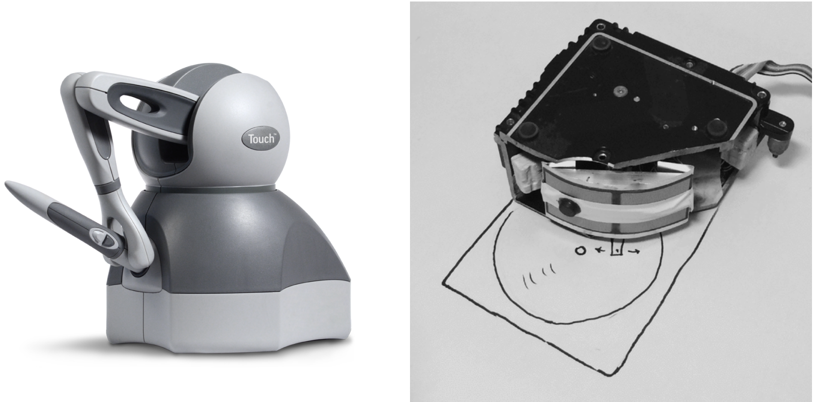
Figure 3.3: Three DMIs that utilise haptic feedback.

**Commercial force-feedback controllers** By far the most popular haptic force-feedback controller which has now become a common device for general research is the Phantom[3] which can be seen in Figure 3.4. There are several variations of this device but in general a stylus is held or a thimble is used to hold a user's finger. This controller has three actuated Degrees Of Freedom (DOF) and can also sense movement in these three orientations. The device functions through the use of small DC motors that can apply torque to the movement of the arm in each direction, dampening and resisting the movement of the users.

**DIY force-feedback controllers** An opensource one-DOF haptic force-feedback controller called the Fire-Fader was created by Berdahl and Kontogeorgakopoulos [25]. This consists of a motorised linear potentiometer slider that utilises a closed-loop feedback control strategy that can be used in conjunction with physical models to simulate the force of interacting with various virtual objects. The Fire-Fader has been used to create many experimental haptic controllers and instruments [25]. The Plank, created by Verplank, Gurevich, and Mathews [363] is an earlier DIY haptic interface that is made from a disk drive that can be seen in Figure 3.4. Again closed-loop feedback control is central to the design of this haptic interface.

**Force-feedback in DMIs** The use of haptic force-feedback transducers in digital musical instruments has also been explored by many [26, 37, 54, 272]. This includes the use of off-the-shelf haptic force-feedback controllers as novel controllers to perform music. Berdahl, Niemeyer, and Smith [26] explored the potential of assisting performers with pitch selection based on variable force from a Phantom controller, finding increased accuracy under simpler feedback condi-

3 https://www.3dsystems.com/haptics-devices/




(a) The Phantom controller (b) The Plank [363]

Figure 3.4: Haptic force-feedback controllers.

tions (detent on in-tune notes); this built on similar work conducted by O'Modhrain [272] that investigated a haptic theremin with tuning assistance using a custom made two-DOF force-feedback device. She concluded that the existence of force-feedback within a DMI can marginally improve the musical task of pitch selection, and that this is dependent on the complexity of the feedback, for example a spring force-feedback setting was less successful than a simpler detents setting.

Similar conclusions have also been reached by Moss and Cunitz [255] in their work on haptic tuning guidance, which used force-feedback to push the musician's finger towards chromatic notes. The tuning guidance paradigm has also been explored by Yoo and Choi [387] with their HapTune device, where two vibrotactile feedback channels are used to guide tuning on a violin. One of the actuation points is above the elbow and the other below: each are actuated to varying degrees depending on the distance from the desired tuned note. In an evaluation of this system they found that it was effective as a chromatic tuner and could be used without interfering with the visual task of score reading [388].

Creating instrument-like haptic force-feedback has been explored with the violin: the vBow [263] aims to replicate a bowed string interaction with a force-feedback mechanism and a physically-derived audio model. Cellomobo [27] is an example of a cello-like instrument with integrated haptic feedback. Bowing a piezo disk attached to a shaker provides haptic feedback to the bowing hand: this approach again integrates both audio and haptic feedback into one feedback loop, and both active and passive touch. Luciani et al. [207] have similarly explored cello interaction with a force-feedback device. Here a physical model controlling both sound and haptic feedback, provided via an ERGOS 2-DOF controller [89], is run on a very responsive DSP platform [206] where the reactivity between input action and force-feedback and sound is 1/44000 ms, and totally synchronous at that rate. Much of the work done at ACROE highlights the importance of responsiveness in the underlying technology used to couple haptic feedback with a physical model of an instrument. Such high performance is required in order to achieve "ergotic sounds": a situation in which the physicality of the interaction is maintained throughout the whole system of the instrument, between hand and ear [207]. They found that the high temporal resolution and deterministic timing of the whole instrument system led to increased playability and fast instrument learning. In individual cases participants remarked on the "strong presence of the string in their hand" which Luciani et al. [207] attest is due to the speed of the whole system.

Similar implementations have been realised in a number of virtual musical instruments that utilise mass-interaction physical modelling and can be played in realtime using haptic controllers [198]. The software platform GENESIS-RT can run haptic loops at 1-10 kHz with acoustic components simulated at 44.1 kHz [198] allowing physical models to be run controlling the haptic behaviour of the instrument in tandem with the virtual sound model. Further development of the ERGOS controller has also resulted in a force-feedback keyboard [89] which can be connected to a real-time digital simulation system allowing the haptic device to be interfaced with a high rate synchronous simulation of the physical model. Similar explorations of interacting with physical modelling synthesis have been covered in the field of sonic interaction design [92, 312].

HAPTIC FEEDBACK AS INFORMATION CHANNEL In a different but related area of research that moves from the instrumental uses of haptic feedback, many researchers have investigated how the haptic sensory modality can be used as an additional information channel during performance. Hayes [133] has experimented with vibrotactile signals delivered to the hands of musicians to offer information about what is happening musically. This was also explored by Schumacher et al. [326] and Giordano and Wanderley [115] by creating a tactile metronome that could deliver timing information to performers in a silent and individual way. Whereas we are focused on applications where feedback is related to musical performance and control, there are many examples of projects that focus on experiencing music through vibration [14, 124, 244], including a personal previous project [153].

Issues relating to research question 1 (How can active tactile feedback be reintroduced into a DMI and what influence does this have on performer experience?) have been explored in the first study of this PhD presented in Chapter 4. The focus of this study is on dynamic vibrotactile feedback and realtime control, and the manner in which this builds upon the research presented above is outlined in

Chapter 4. I also introduce a novel signal processing technique for providing audio-related vibrotactile feedback that is based on the acoustic phenomenon of 'beating' that occurs between two closely tuned notes.

## 3.2 THE EMOTIONAL CAPACITY OF TOUCH

The emotional capacity of touch is central to the field of Tangible User Interfaces (TUI) which emerged from research in HCI in the mid 1990s [151]. Many of the foundational ideas of this field have grown from what Fitzmaurice, Ishii, and Buxton [88] defined as Graspable User Interfaces: interfaces that allow direct manipulation and control of a digital object via a tightly coupled physical handle. This has led to much interest in how physicality, movement and touch can be exploited when interacting with a computer. Following on from this, Ullmer and Ishii [358] defined the central characteristic of tangible interfaces as the coupling of physical representations to underlying digital information and computational models. The field is representative of third wave HCI and was founded with a strong basis in embodied cognition and an initial stance against the dominant GUI [131]. By focusing on the connection of the physical world we inhabit with the digital world of the computer, the field argues for a rediscovery of the 'rich physical aesthetics of manual interaction' [328].

### 3.2.1 *Tangible and physical computing*

What makes a tangible interface different from a haptic interface? Tangibility itself is a rich and complex term whose definition is still a work in progress across many different fields. Tangibility is the property of an entity to be accessible to the sense of touch: physical contact of some description is central [145]. As Cadoz et al. [49] note, in a broader sense a thing is tangible if it is real and not only imaginary, if it is defined and not vague. A number of associated terms come packaged with this word: reality, materiality, objectivity, presence, concreteness can all act as synonyms although each with their own nuanced differences in meaning. Almost all of these terms are used as qualifications of whether an object or entity is genuine or not – this is related to the notion that the eyes and ears can be deceived but if you can touch it with your own hands then you know it's real [286]. Two further extensions of the idea of tangibility which distinguish it from other senses are 'immediacy' and 'manipulation' [49]. Immediacy refers to the fact that the sense of touch relies on direct physical contact, giving us a direct and *intimate* experience of an object that the other senses can't provide. Touching is our primary means of effecting change on objects in our environment, and manipulation refers to the bi-directional nature of touching: it is at once an input and an output act. Tangible computing, then, capitalises on the unique characteristics of physical manipulation rather than putting them in service of the computer, and brings to the forefront the physical nature of interaction in all its richness. Dourish highlights the importance of attaching more meaning to the physical aspects of the interfaces we use to control computers.

> Research into tangible computing has taken a step back and realised that, while we currently interact with computers through physical objects (such as keyboards, mice and displays), we can better exploit our natural skills if we focus on interacting with the physical objects themselves. The physical objects no longer stand as proxies for purely computational entities like cursors and insertion points, but can begin to take on a more direct role in the interaction. [75, p. 7]

For tangible interfaces, besides the exploitation of the affordances and constraints of the physical object, it is the creation of rich haptic experiences that has thus been one of the driving ideas [328]. To return to the question at the beginning of the section, whereas haptic interfaces aim to physically stimulate the sense of touch of a user, tangible interfaces can be regarded as taking an expanded view of the different elements that can contribute to haptic experience, relying more on metaphors of 'natural' interaction within interface design.

#### 3.2.1.1 *Tangible interaction*

A classic example that is often cited as a point of inspiration for the development of tangible interfaces [151] is the Marble Answering Machine which is a concept sketch created by product designer Durrell Bishop [1]. In the Marble Answering Machine incoming calls are represented by coloured marbles that roll into a bowl when a call is missed. To playback the message that has been left the marbles are placed on an indentation, and can also be placed on an indentation on the phone to call back the person who left the message. This playful design study is popular within the field due its clear reliance on physical affordances and everyday knowledge. Bishop's work assigns new meaning to the marbles, transforming them into containers for data and references to other objects in a network: he use known objects as clear references to the aesthetics of new electronic products [328].

In the short period of time that this field has existed it has produced a plethora of novel interfaces. Sandscapes [152] is an interface that allows museum visitors to model different geographical landscapes by manipulating sand or clay: through level sensing and projection mapping a dynamic ecosystem is created where the visitor to the museum shapes the lay of the land with their hands. DataSpoon [392] is a tangible medical device that consists of an instrumented spoon designed to monitor the movement kinematics of children with motor disorders when self-feeding. It provide caregivers with information from which they can assess the recovery of the child. ReFlex [346] is a bendable mobile device that allows users to skim through pages of an e-book by bending the interface in different ways. While the initial definitions of a TUI were generally focused on the representation and transmission of information, a growing number of projects instead focus upon tangible interactions in a wider sense; that is human action, control, expression and social dynamics around a system [145]. These are factors that are essential to musical interaction.

### 3.2.2 *Tangible musical interfaces*

There are many examples of tangible interfaces that have been created for musical purposes: Martin Kaltenbrunner has created an online gallery of projects dating from 2000 to around 2009 that showcases tangible musical interfaces[4]. Kaltenbrunner organises the interfaces into the following non-exclusive categories: *tangibles, blocks, tokens, artefacts, toys, touch, controller, malleable*. This gives an indication of the scope and diversity in approaches but also the binding traits of the musical projects he present as 'tangible'. A common theme in the instruments presented is table-top implementation with back projection with tracked tangible tokens that move across the table's surface to control parameters of the sound. The most famous example of this (and of tangible user interfaces in general) is the reacTable [168], a musical interface for collaborative control of sound synthesis processes. The system allows users to arrange specialised physical tokens on an active surface. Their configuration controls the parameters of filters, sound generators, and sequencers, each associated with the physical objects.

Other notable examples of tangible musical interfaces are BeatBearing [23] by Bennett and O'Modhrain which takes Bishop's idea into the realm of musical control by using a series of ball bearings that users can arrange on a grid to control rhythmic patterns. AudioCubes [322] are an example of a block based TUI where multiple individual blocks with individual behaviours join together into a musical system with lights synchronised to sound. Various educational or playful interfaces aimed at different groups of performers have also come from the field, for example the TouchTone [28].

Increasing importance is being given to experimental musical instruments and instrument building practices, both as inspiration for work in tangible computing and as a history lesson, demonstrating

4 https://modin.yuri.at/tangibles/

how early research in musical interactions with electronics and computers can foreshadow developments in this field. For example Andersen and Ward [5] have recently presented a description of the Crackle Exhibition with took place in 1975 and was developed by Michel Waisvisz in collaboration with STudio for Electro-Instrumental Music (STEIM) an independent electronic music research centre based in Amsterdam, Netherlands. The exhibition consisted of 20 touchable sound-producing objects. Andersen and Ward highlight how this experimentation relates to current state-of-the-art developments in Tangible, Embedded and Embodied Interfaces (TEI), and in some cases surpasses it. Shared research between NIME and TEI is increasingly happening, and the research presented in this thesis is representative of such a cross-over with elements of this thesis being presented at TEI 2016 [155], TEI 2017 [156], NIME 2016 [239] and NIME 2018 [159], and part of the work this thesis aims to do is to bridge discourses in both fields.

It is interesting here to consider what the difference is between tangible musical interfaces and musical instruments. What are the unique characteristics of a musical instrument that make it instrument-like? In his recent exposition of organology and the place of DMIs and other musical devices within it, Magnusson [217] builds upon the definitions of *instrumentality* put forward by Hardjowirogo [128]. For her, instrument identity depends on seven potential criteria: (1) sound production, (2) intention/purpose, (3) learnability/virtuosity, (4) playability/control/immediacy/agency/interaction, (5) expressivity/effort/corporeality, (6) 'immaterial features'/cultural embeddedness, (7) audience perception/liveness.

From the point of view of musical instruments, there are some trends in tangible interfaces design that could be considered as shortcomings. There is often a strong reliance on visual feedback for things like state changes, notification of task completion or task monitoring. The majority of tangible musical instruments tend to be musical systems which are based on higher-level interaction with musical structures, as in the case of a sequencer or Digital Audio Workstation (DAW), rather than performance instruments that rely on closely coupled action and perception. They also mainly target novice users and focus on collaborative music making spaces where there is a short interaction time and a mix of musical abilities, and perhaps for this reason the low barrier to interaction, offered by that something like a sequencer for instance, is well suited.

### 3.2.3 *The challenges of DMI design*

Musical instruments pose a series of challenges that are not usually addressed in tangible interface design: that of realtime control, the close coupling of action and sound, and the goal of expressive interaction. Musical instruments exemplify a specific and specialised case of human-computer interaction: they serve as important examples of tools that foster a complex and meaningful exchange between user and device. They are also tools in which the physical or tangible elements of their design are of particular importance. This is due to the orientation of instrumental practice towards embodied activity [196] and highly specialised sensorimotor skill [213], paired with the complex cultural signification of the device itself and the music it produces [215]. Tzanetanis et al. have even argued that musical instruments anticipate aspects of rich multimodal and embodied interaction by providing "excellent examples of sensorially rich and temporally detailed human-machine interaction" [357, p. 1119].

As discussed previously, sound production and the translation of movement are two factors that are intrinsically coupled in musical instruments. Schloss [324] proposes that part of the reason people attend musical performances is to witness skilled players doing something they cannot do themselves. When action is far-removed from sound production, appreciation of the skill of the performer becomes difficult [97]. Wessel observes that "[a]s computers begin to populate the musical stage they are most often found before performers who manipulate the keyboard and mouse in a manner all too reminiscent of office work. This is certainly not a situation that invites the development of musical virtuosity nor enthusiasm on the part of the audience" [373, p. 93]. In the following section we shall review some of the varied approaches to remedying this situation and making interaction with DMIs more physical.

#### 3.2.3.1 *Movement transducers*

The intrinsic connection of movement and musical meaning that we discussed in Section 2.2.1 has deep-reaching implications for the way we conceptualise and design instruments. In his discussion of organology and the challenges that digital and electronic instruments pose to the categorisation of such artefacts, Kvifte points out that the linkage between performance technique and the design and acoustic qualities of an instrument is so strong that it does not make sense to talk about one without the other [188]. From another perspective, the ethnomusicologist John Baily has worked extensively on the connection between the physical layout of an instrument and the body movements that this allows. This research has led him to describe acoustic instruments as *movement transducers*:

> 'Acoustic music' is the product of human movement processes and embodies aspects of the human sensori-motor system, which to some extent and in various ways shape the structure of the sonic product. Musical instruments

> are like machines with which human sensori-motor systems interact. The instrument itself has an 'active surface' in relation to which the body moves. A musical instrument is a type of transducer, converting patterns of the body movement into patterns of sound. [15, pp. 123-124]

Acoustic instruments are here understood as amplifiers of movement patterns which transform the kinetic energy of a musician's movement against the 'active surface' of the instrument into acoustic energy, perturbations of air pressure that bear the trace of the action that created them. This seems undeniable in the case of acoustic instruments where there is a direct coupling of action and sound [163], what Peters's terms 'real' touch, as opposed to 'apparent' (simulated), or 'absent' (deliberately avoided) [289]. As the mechanism of an instrument takes more agency over sound production, the action of the performer and sound that the instrument produces become increasingly decoupled.

In the case of DMIs, the coupling of action and sound no longer relies on mechanical translation or transduction of kinetic energy into acoustic energy; rather, sound production is the result of a designed chain of mediation. Magnusson reminds us that "[d]igital instrument makers [...] get nothing for free, unlike makers of acoustic instruments who receive the gift of sonic timbre from the physical properties of the materials they work with" [215, p. 44]. Here is a typical chain of mediation: to begin sensors convert movement into electrical signals, these electronic signals are converted to digital signals which are then mapped to parameters of a sound generator, and finally the signal from the digital sound generator is converted into an electronic signal that drives a sound transducer to create soundwaves. This complex chain of mediation has led Nijs to describe digital musical instruments as mediators between musical gesture and musical sound [264], and given rise to discussion of the degrees of action-sound separation [163] that relates to how closely action is connected to sound.

## 3.3 CONTROL INTIMACY

*Control Intimacy* is a term first introduced by Moore when describing the musical dysfunctions of the communication protocol MIDI (Musical Instrument Digital Interface), which at the time of writing his article (1988) was growing in usage and quickly becoming the standard way of interfacing with digital musical systems and instruments [252]. Intimacy has since become a central criterion of DMI design and has been been expanded by many in the field [83, 166, 251, 375]. Almost thirty years later MIDI is still by far the most commonly used protocol in digital musical systems despite the advancements given by Open Sound Control (OSC) [383]. In the same article Moore discusses MIDI from a musical point of view, focusing on how it is used to capture musical performance and act as a digital representation of musical control processes. He states:

> [f]or subtle musical control to be possible, an instrument must respond in consistent ways that are well matched to the psychophysiological capabilities of highly trained performers. The performer must receive both aural and tactile feedback from a musical instrument in a consistent way – otherwise the instrumentalist has no hope of learning how to perform on it in a musical way. [252, p. 21]

Moore identifies traditional acoustic instruments as possessing this important quality of 'intimacy'. The voice is given as an example of the most intimate of musical instruments: when singing we use vocalic control that is innate and informed by speech as much as music. Similarly the violin, sitar, and flute are given as examples of intimate instruments as they allow the micro-gestural movements of the performer to create a wide range of stylistic, expressive and affective variation in the control of musical sound. With MIDI tiny variations in performance, whether temporal or spatial, are not reflected in the sound, preventing these gestures from having any effect on the music that can be controlled by the performer. For Moore MIDI represents a 'degradation of control intimacy': at once this is about the imprecision of timing in MIDI and the lack of resolution and compression it applies to captured musical gestures [252].

Control intimacy is a useful concept for DMI design as it can be observed both qualitatively and quantitatively [166]. Qualitatively it relates to how an instrument translates the gestures of a performer into sound, and a performer's impression of how it responds to their movements. Quantitatively it relates to many design decisions that concern the resolution and sampling of movement by the digital system of the instrument. *Ergoticity* in DMI design is a related concept that developed from the work of Cadoz and Wanderley [47, 48]. Cadoz suggests that an essential property of instrumental interaction is the preservation of energy through both digital and physical components of a system: all signals produced by the system should correspond to the amount and shape of energy fed into the system so as to provide a more natural form of interaction.

### 3.3.1 *Spatial aspects: sensor resolution*

Wessel and Wright expand on Moore's notion of control intimacy and its relationship to virtuosity [374]. They state that low latency responsiveness is essential for control intimacy, and in particular low variation of latency (jitter), but also highlight the importance of the fidelity with which a gesture is captured by a computer. They argue that many musical gestures are continuous functions of time and should be treated as such, rather than being split into discrete events and triggers with velocity, as happens with MIDI. They propose OSC as an alternative communication protocol for DMIs [383] that moves away from the 'control signal' based approach of MIDI and towards a communication protocol where everything can be treated as a continuous high-definition signal. The importance of audio-rate sensing of gestures and its direct relationship to control intimacy is also put forward in relation to CNMAT's connectivity processor [13, 375], which has the ability of reading continuous gestures into the computer in a manner that is very tightly synchronised with the audio sample stream, an approach that is shared with Bela [240], the platform that has been used to build all the instruments in this research. This also mirrors the principles behind the high performance DSP system used with ERGOS and built at ACROE [89].

Writing about intimate experiences of musical control and audio-rate sensing, Wessel [373] urges the reader to try placing a Force-Sensing Resistor (FSR) on their finger tip that is sampled at audio rate by the computer, and used to scale the amplitude of a sine tone. With just this primitive but direct set-up a remarkable amount of timbral and temporal variation can be achieved by tapping in different ways and on differently textured surfaces. The variety and expressiveness of this simple instrument is down to the resolution of the sensing (audio quality bit rate instead of the 128 steps of MIDI) and the tight feedback loop. In Chapter 6 I present a study that evaluated the effect of variable levels of control intimacy on instrumental performance.

### 3.3.2 *Temporal aspects: action-sound synchrony*

The temporal aspects of control intimacy concern the alignment of action and sound and their temporal coupling. Latency has been identified as a barrier to virtuosic engagement by obstructing a fluent interaction with the instrument [218, 273, 382]. This is because it prevents the kind of responsive and immediate interaction that is possible with acoustic instruments: tools that foster a relationship between gesture and sound that is both intuitive and complex [73]. Wright states that a few milliseconds of latency and jitter can make the difference between a responsive, expressive, satisfying real-time computer music instrument and a rhythm-impeding frustration [382]. Due to the complexity of the sensorimotor control that a musician has over an instrument and the high demands of musical performance, DMI design is a good testing ground for understanding the effect of latency and jitter in HCI more broadly, complementing research done in relation to musical disruption caused by DAF. In this thesis I propose that the effects of latency are fundamentally tied to tangible aspects of a DMI's design, to what could be described as its 'feel' dimension.

A small number of studies have investigated the impact of latency on DMIs, both in terms of latency perception and of its influence on performance. There still remain few studies that investigate latency below the threshold of simultaneity perception, and this is a gap that the research presented in this thesis addresses. The studies that do exist in this area have focused on larger delays. Instruments with continuous non-physical control, for example, have been shown to be less sensitive to latency than those that involve direct contact: for a theremin, which is played with in-air gestures of the hands that never touch the instrument, the Just Noticeable Difference (JND) for latency perception was shown to be around 20-30ms with latencies as high as 100ms going undetected during the performance of slow passages [219]. Dahl and Bresin, in a study with in-air percussive digital musical instruments without tactile feedback, found that a latency of 40ms negatively impacted timing accuracy, but that up to around 55ms performers were able to compensate for the latency by increasing their anticipation (moving their strike earlier) when latency was gradually introduced [66]. Larsen and Knoche [191] have investigated the influence of DAF of 73ms and 250ms on the strumming performance of both non-musicians and musicians. They found that when the DAF was matched to a subdivision of the target metronome tempo (in this case 250ms delay with a metronome of 120bpm (IOI of 500ms)) the impact on timing was minimal for musicians but there was a substantial decline in the timing performance of non-musicians.

In Chapter 5 I present an experiment which addresses these issues, aiming to discover some of the fundamental characteristics of action-sound latency in DMIs that have remained under explored up to this point. Although much is known about larger delays, and there is general agreement that latency in digital systems has a negative effect, few studies have been conducted that investigate how latency relates to the perceived quality and tangibility of an instrument and how this delay impacts upon rhythmic performance. I was also interested in the impact of highly specialised rhythmic training on latency perception, and so decided to work with both amateur musicians (with varying degrees of musical accomplishment) and professional percussionists (with a high degree of speciality in rhythmic performance).

### 3.3.3 *Transparency*

In HCI the notion of *transparency* has been central to the discourse around interface design and evaluation for many years. Transparency refers to a person's ability to use an interface with no or little conscious effort directed at the interface itself [35]. Achieving transparency in an interface is dependent on a good *coupling* between the digital and physical parts of a system. Fels, Gadd, and Mulder [84] build on this definition and define transparency as the quality of a mapping representative of the 'psychophysiological distance' between the input and output of an instrument in the minds of the player and the audience. This notion contains clear echoes of Moore's control intimacy: full transparency for a performer means that a device's output exactly matches a performer's expectation and control [84].

Fels posits that the end goal of musical instrument design should be for the player to have a high degree of intimacy, to the extent that they embody the instrument: at this point the instrument behaves like an extension of their body – there is a *transparent* relationship between control and sound [83]. This allows intent and expression to flow from the player, through the instrument, and to the sound, creating music. Nijs, Lesaffre, and Leman [264] have refined this concept in terms of musical embodiment. At a certain level of expertise the instrument can become transparent and act as a mediator between musical intention and musical result. At this point the musician stops considering an instrument's individual operations, and rather focuses on higher level musical concepts, phrasing, articulation and spontaneous corporeal expression.

The notion of transparency incorporates key concepts from phenomenology regarding tool use, for example Heidegger's notions of 'ready-to-hand' and 'present-at-hand'. Through skilled use, a tool can 'disappear' into its function and becomes transparent to the user. When a tool malfunctions or has a bad coupling, it increases the chance of it becoming apparent to the user: a Heideggerian 'breakdown' happens transitioning the interface from 'ready-to-hand' to 'present-at-hand' [136]. Tools and technologies extend our potential for action as it emerges from our interactions in the physical world. Merleau-Ponty's famous example of the blind man's stick becoming an extension of the lived-body illustrates this point well [246]. In the case of this blind man, Merleau-Ponty explains, the perceived world does not begin at the point where the hand holds the stick, but rather at the end of the stick: the stick is transparent to the blind mind. It is important to note here that transparency is a characteristic of the interaction between a tool and a user, and cannot be assigned directly to a tool or user themselves.

Bolter and Gromala [35] remind us that transparency is not necessarily the only end goal when designing interactive systems and that there is in fact great benefit in the un-transparent tool. They suggest that interfaces should oscillate between transparency and reflection, windows and mirrors, allowing the user to "step back and contemplate" while also allowing the medium that underwrites the project to not disappear [35]. Their statement is a reaction to Norman's dictum from *The Invisible Computer* [268] which talks about the computer becoming invisible and stopping to be present as a computer; rather, it begins to be perceived as the function that it fulfils (e.g. phone, calendar, camera, music player). Norman compares this to the way that the motor is forgotten about in the vacuum cleaner or blender.

With new musical instruments designers often want the technological mechanism of the instrument to disappear: instead of conceptualising a new instrument as a sensor that when tapped upon produces a musical tone from a computer in response, the goal is for it to be considered an instrument fit for music-making, and which encourages music-making. However, a way in which musical instruments complicate the notion of transparency is by being examples of tools that are not designed to be immediately understandable and easily learned. Coyne, Parker, and Rebelo [63] describe the cello as an instrument of discipline rather than a harmonious continuation of human agency, and argue that an ergonomically neutral cello would strip the instrument of its potency: certain tools, like musical instruments, are characterised by a resistance to the *seamless interface* [151]. They posit that instead of the free and direct flow of gestures being an expressive act, it is in the contact, resistance and labour of performance that musical meaning lies [63].

This relates to the problems we identified with the notion of 'natural mappings', mentioned above. Hornecker [144] warns that this naturalness should not be taken for granted or treated as easy to achieve. When designing a system with a set of action affordances it is impossible to account for every action a performer will use with that system: affordances shift with the user and the environment, leaving all devices open to appropriation. The *naturalness* and *intuitiveness* of tangible interaction is one of its primary selling points, but the difficulty and amount of effort it actually takes to fulfil this promise is slowly being realised across the field [144]. Different approaches are required and this is because computer systems, by their very nature, are not like the real world. Natural interaction seems a 'holy grail' that is unattainable [269]: aside from digital verisimilitude (the case of the *uncanny valley* of instrumental control, where a DMI is almost indistinguishable from an acoustic instrument in its behaviour) there are other aspects of instrument design that can have great influence on the tangibility of an instrument.

## 3.4 TANGIBLE GUIDES AND CONTROL METAPHORS

Essl and O'Modhrain [80] introduce the enactive approach to instrument design as a means of preserving the coupling between a musician's sensorimotor capabilities and their understanding of the physical properties of tangible objects. By utilising 'familiar sensorimotor experiences' mapping can become less of a digital question and more of a physical one.

For Essl and O'Modhrain instrument design should build upon tacit knowledge acquired from interacting with objects in the real world: for example, when one plays with rocks, shuffles leaves, or moves one's hand through water, there is in each case a tight coupling between actions performed and the tactile and acoustic responses to these actions. This knowledge can be applied to other cases of interaction, where similar actions can be expected to maintain some of the character of previously experienced interactions [312], leading to what they term physically-inspired mappings between action and sound.

One of the best examples of this is PebbleBox [274], a granular synthesis-based instrument where the main interface is a box filled with rocks that the performer moves and collides against one another. There is a microphone on the base of the box that captures the sounds of the rocks colliding and scraping against one another. The sound captured by the microphone is used to drive a granular synthesis engine which works with sound materials that share characteristics with the colliding rocks: water droplets, crackling fire, ice cubes. When playing the instrument, experience tells the performer that colliding pebbles will have an auditory and tactile effect with particular forces and frictions. By taking the physical experience of a particular type of interaction and applying it to another type of sonic interaction that is similar in terms of physical energy, PebbleBox manages to double the interaction metaphor through both the physicality of the instrument and the resultant sound.

Similar themes have been explored at length in research into sonic interaction design which, with a focus on everyday environmental sounds rather than necessarily musical ones, investigates the ecological validity of the linkage between an action and a sound [71]. This field's standpoint is that through modelling sound sources according to their physical behaviour it is possible to define a natural mapping between human gestures and the control parameters of the sound model, this way providing physical consistency between action and sound [92]. In both these approaches the notion of naturalness is less dependent on digital verisimilitude and more on the tacit knowledge we have from interacting with the environment and tapping into these reserves.

### 3.4.1 *Navigating a musical instrument*

A good deal of Chapter 2 focused on sensorimotor skill acquisition and the structuring of an internal model through extensive interaction with an instrument. When we interact with a computer we form a similar model that relates the state of the computer to our position in the interaction. In his *Interaction Design Sketchbook* Bill Verplank

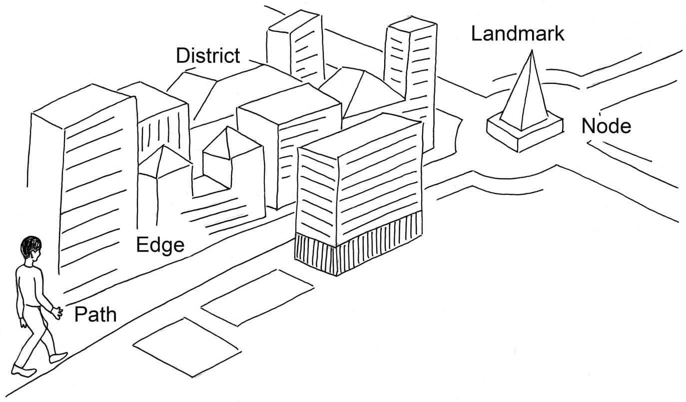



Figure 3.5: Elements of the city, adapted from Lynch [208].

[362] proposes that the work of urban planner Kevin Lynch [208] can provide a useful framework when considering visual window-based programme design. Lynch writes that the best urban planning supports not only efficient routes through a city, but mental maps, a quality that he terms the "imageability" of a city. Lynch asked a number of citizens of various American cities to describe routes and sketch maps of their city. From this study he developed the following classification of five elements: *Landmark*, *District*, *Edge* (between districts), *Path* and *Node* (where two paths intersect). He found that the cities that citizens were able to sketch best shared a certain relationship between these elements, for example paths along edges so that districts are distinct, or landmarks at nodes so they can be used as reference points during navigation.

The usefulness of this method to screen-based applications is clear where navigation is reflective, complex and there are often hierarchical page structures and, with bad design, places where you can get stuck, get lost, and not return from. The term 'imageability' seems on the surface to refer to the visual arrangement of a city, but it is also about the metaphors we employ to understand everyday perception, which are far from unisensory. This work echoes theories of spatial schematas from second wave HCI which investigated how people learn to touch-type and found a series of movements that served as the anchors and building blocks of motor learning [335].

It is possible to relate Lynch's elements to the musical topography of an instrument, in particular to how the design of an instrument arranges them and how their formations create musical topographies, as discussed in Section 2.1.4.2 in relation to Rojas [313] and Sudnow [348]. Alongside the physical layout of an instrument, it is the instrument's reflexive nature, it's active multi-modal response to our input gestures, that leads to the formation of such topographies. Rebelo posits that expressive engagement with a musical instrument happens through "micro-level (haptic) variation, which suggests orientation and negotiation that is articulated step by step, at a local level" [304, p. 31]. This engagement is guided by both static factors (material, weight, arrangement of keys, strings or frets) and dynamic factors (how it responds to energy put in by the performer) [250].

The arrangement of an instrument could be compared to a *tactile map*, a physical representation of a geographical area made available to the sense of touch, usually designed with blind people in mind. In these maps certain landmarks are blown out of proportion, and simplifications are made to areas of the city to encourage an understanding of the city's main arteries, its global structure and important sites. In many ways these maps could work as a metaphorical comparison for an instrument, but in the case of an instrument it's more complex because the haptic map of an instrument is dynamic in its behaviour, and is designed not just to provide information via its layout, but to translate action to sound when movement aligns with the correct nodes, districts or landmarks in its design.

In terms of 'imageability' then, we can consider which factors in an instrument's design behave as landmarks and nodes that allow us to build this internal model of instrument behaviour after interacting with it over an extended period of time. The physical supports of certain musical gestures (for example strings, keys, valves, buttons, slides, bows, pedals) provide the hands with passive haptic guides that instruct them as to how they should be arranged and tell them how a note should correctly be produced on the instrument. If we also consider the dynamic behaviour of each of these elements then we can appreciate the trajectories of musical movement that each of these categories of musical input device allows: each has a defined series of paths of gestures that are able to produce sound. It's perhaps in part through learning these pathways, and discovering 'efficient' means of producing sound that a musical culture develops with an instrument.

### 3.4.2 *Static factors*

MATERIAL CHARACTERISTICS In the last few years the topic of materials has surfaced as a major theme within the research field of interaction design [85, 376]. The materials that an instrument is constructed from, whether that be wood, metal, rubber, plastic, ceramic, glass or stone, already contain a great deal of information within them: different materials, in the characteristics they display, have different action affordances[5]. These affordances can be classified into

5 Sensors themselves can act as their own type category of material, one that contains an expectation of certain types of movement and the tracking of gestures in certain ways [298].

the following categories: texture (rough or smooth), density (hard or spoungy), temperature (hot or cold), perceived strength (durable or fragile) [100]. As mentioned in Chapter 2, these haptic qualities of a material are bound to different types of exploratory hand movement: there are certain set patterns of movement that allows us to test out these qualities [195].

The characteristics of a material are often exposed through their sonic behaviour when interacting with them. This idea was explored in an instrument created by Merrill, Raffle, and Aimi [247] that consisted of a pallet knife with a piezo vibration sensor attached. The signal captured by the piezo was fed into a convolution algorithm that maintained acoustic characteristics of textures and different materials that the knife was scraped over. In an installation context the knife was presented alongside various materials including bathroom tiles, sheep's wool, broom bristles with varying stiffness, artificial turf, aquarium pebbles, shag carpeting, metal screen, and wicker curtain pieces. This study echoes work we discussed in relation to pseudo-haptics and the power of multisensory effects in Section 2.3.4.

WEIGHT AND PERCEIVED QUALITY In consumer research the linkage between, amongst other tangible qualities, the weight of an object and its perceived quality has been long known and utilised in "tactile branding" [338]. Some common examples on this can be seen in the weight and texture of luxury crisp packets, or the "soft touch" resins, that provide a particularly soft and pleasurable feeling when held in the hand, increasingly used in beauty products instead of hard plastic. In terms of weight itself, in an experiment conducted by Piqueras-Fiszman and Spence [299] a correlation between the weight of a bottle of wine and its perceived expense was found. Similarly, the weight of cutlery as well as its size and shape, has been shown to affect the taste of food [129]: yoghurt was perceived as denser and more expensive when tasted from a lighter plastic spoon as compared to an artificially weighted spoon.

Klatzky and Peck [180] investigate the visual aspects of a product's design that invite touch, identifying what they call the 'touch-ability' of an object: visual aspects that elicit touch. They found some dependencies of the structural properties of the objects on touch-ability but there was also great difference between ratings. Other consumer research suggests that touching a product can increase the perceived value of an object [288], encourage consumers to make more unplanned purchases of an object they touch [287], and enhance feelings of "ownership" [288].

ARRANGEMENT AND LAYOUT Jordà [167] identifies a issue that effects DMI design. With acoustic instruments the physical form of the 'controls' typically arises directly from the acoustic process controlling that instrument. Their shape is therefore communicative of their role in a way that those on generalised DMI controllers are not.

Many of the physical aspects of DMIs are defined by the demands of the sensing strategy, but studies into the influence of physical form in instrument design have demonstrated how simple modifications to the material, form and dynamic behaviour of an input device can have great implications for how a performer interacts with an instrument [24, 161]. In the space of musical interfaces this has been explored by Jense and Eggen [161] who created numerous physical variations of a potentiometer input device. The methodology employed in this design exercise is noteworthy for its playful exploration of the dynamic behaviour of a fundamental input device. Six knobs that varied in size, form, material and dynamic behaviour were created to explore the expressive affordances of such a simple input device, and to see how these variations had an impact on user experience.

### 3.4.3 *Dynamic factors*

INPUT MODALITY DMI designers are responsible for the material characteristics of the instruments they build, both the static factors and the dynamic behaviour that provides the performer with feedback. What I term as 'input modality' aims to account for more than simply sensor choice: it is the sensing strategy and the particular type of interaction that the physical behaviour of the sensing mechanism encourages. It has been noted that there are certain patterns of movement that are better suited to particular sensors [229]. Alongside the sensor itself, and the particular type of movement it captures, the housing of the sensor and physical supports around the sensor also condition how the performer plays the instrument.

In a keynote given by O'Modhrain at Eurohaptics 2016[6] she spoke of how the focus in musical instrument design should be on the transferral of energy in a system, echoing notions of the ergotic DMI design [207] (the maintenance of energy through all aspects of a DMIs system) and her earlier work on enactive musical interfaces [80] (utilising familiar sensorimotor control). In terms of energy transferral there is the vibration of the instrument happening at audio rate, but there is also a transfer of energy that happens at a much slower rate, what she terms the interactional rate. Striking an instrument injects energy into its system but also has an energetic effect on our control gesture: the action has a reaction which colours how we interact with the instrument. This is summarised into acoustic rate behaviour (how the instrument vibrates) and interactional rate behaviour (how the instrument pushes back).

6 http://www.eurohaptics2016.org/

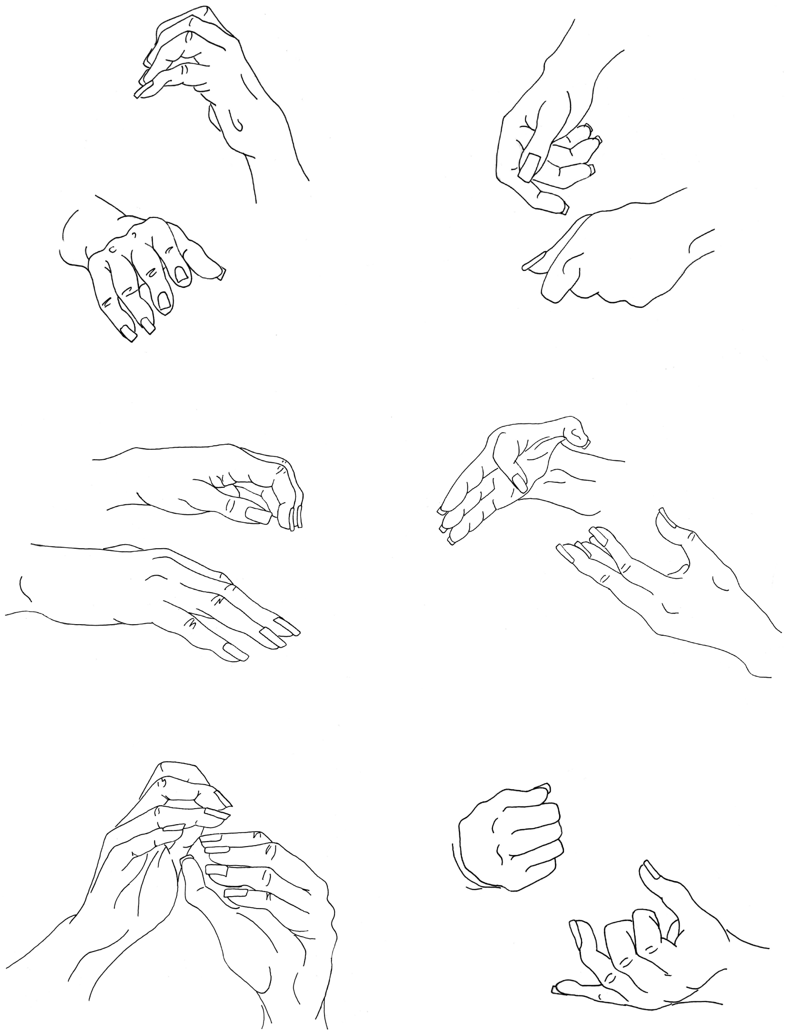


Figure 3.6: The positioning of hands while playing various instruments.

It can be useful to break down haptic feedback into the audio and interactional rate responses of the instrument: how it is acoustically excited, but also its interactional response. For the interactional rate we see the biomechanics of the player and the mechanics of the instrument becoming coupled in the moment of excitation. The player has to manage the two parts of this coupled system, their body and the material behaviour of the instrument as they act on it.

### 3.4.4 *Cultural factors*

INSTRUMENT FORM In his theorisation of tangible user interfaces Horn points out that the evocation of cultural forms can tap into users' existing cognitive, physical and emotional resources, activating existing forms of social activity [143]. Musical instruments, and the physical forms they take, are some of the most potent artefacts for how they evoke existing cultural practices. The topography of the instrument, for Rojas, is not only dependent on the layout and mechanical response of the instrument: it also pertains to an expressive musical sphere which emerges through knowledge of the sound the instrument creates, knowledge of the effective gestures used to control the instrument, and knowledge of the response of the materials (in both their physicality and their cultural significance) from which the instrument is constructed [313].

Mastering a musical instrument, then, is not just the development of a sensorimotor skill, or the efficiency of sound production, it is also a cultural practice aimed at creating a cultural product (music). Small introduces the notion of 'musicking' to capture the way in which any form of music demands its own set of material performance and listening practices [334]. For Small music is defined as something that we 'do': musicking highlights music as an activity that is surrounded by sets of behavioural laws that are offered to both performers and perceivers. Much ethnomusicology is dedicated to the study of the cultural specificity and importance of the laws that surround musical practices. A particularly relevant example comes from Bates:

> Much of the power, mystique, and allure of musical instruments [...] is inextricable from the myriad situations where instruments are entangled in webs of complex relationships between humans and objects, between humans and humans, and between objects and other objects. Even the same instrument, in different sociohistorical contexts, may be implicated in categorically different kinds of relations. I thus am arguing for the study of the *social life of musical instruments*. [18, p. 364]

In a similar vein, Bijsterveld and Schulp [29] discuss tensions surrounding innovation in traditional instruments through conversations with instrument manufacturers who have employed innovative approaches in their designs. A key example is the Pellegrina, a radical redesign of the traditional viola which allows performers to access higher positions on the neck without risk of injury or discomfort. The Pellegrina's striking, asymmetrical shape led to initial shock amongst other orchestra players and garnered substantial press attention. According to Rivinus, the designer of the Pellegrina, "if it had only sounded new, reporters wouldn't have been nearly as interested" [29, p. 666]. Despite its eventual popularity, the initial negative response to the Pellegrina in classical orchestras is symptomatic of a musical culture in which expectations and requirements on instrumentation and performance are rigidly upheld.

## 3.5 EVALUATING DIGITAL MUSICAL INSTRUMENTS

How to judge an instrument and evaluate the success of its design is a core issue in areas of research that deal with DMIs. The *effectiveness* of a musical instrument can take many forms and vary greatly depending on the stakeholders and their associated expectations [273]. A performer, for example, may judge an instrument on its ability to allow them to play compelling music live: the success or failure of this performance is often tempered by the response of an audience and by whether they found the music engaging. There has been an increase in the application of user-experience evaluation techniques from HCI to DMIs [44], but the complexity of the creative context that DMIs exist within poses a series of challenges to such techniques. The centrality of performance to the evaluation of a DMI indicates that a more expansive form of evaluation is necessary, one that accounts for the many stakeholders at play and their cultural context.

In addition to performers and audience O'Modhrain [273] identifies composers, instrument builders, component manufacturers and consumers as stakeholders in an instrument's design, each with a different idea of what 'evaluation' means. The multitude of different perspectives on DMI design suggests that various evaluation techniques and approaches may be required. There have been many different methodological approaches to this challenge, from frameworks [145, 221, 362] and guidelines [61, 369] to models [148, 189] and taxonomies [218, 280]. All share the desire to analyse, compare and contrast different pre-existing instruments and interfaces. As well as looking back and taking stock, these approaches also want to inspire future design by encouraging designers to work from a particular shared conceptual base.

In terms of actual evaluation techniques, due to the complexity I discussed above, it is clear that quantitative techniques alone would be insufficient to represent the various perspectives at play. Instead a balanced mix of quantitative and qualitative techniques are often employed. In what follows I shall review the various approaches that have been taken in the field to assess an instrument's *functionality*, its *quality* and the resultant *performer experience*.

### 3.5.1 *Functionality*

As technological devices DMIs often offer novel musical functions to a performer through their affordances. One of the most basic forms of analysis in the field therefore relates to testing the function of a DMI: a design's practicality, performance, stability and sturdiness. Some instructive examples of this come from papers that introduce a novel technology or implementation strategy for DMIs, providing extensive information on the performance of this strategy in comparison to others, for example in regards to latency performance [239], sensing resolution [166] or behaviour of a communication protocol [383]. This type of evaluation usually relates to more fundamental components of DMIs, with the analysis happening without any kind of performance being necessary: this relates directly to the kind of device testing common in engineering and product design.

#### 3.5.1.1 *Definition through control strategy*

The functionality of a DMI is commonly represented as the relationship between a series of inputs and a series of outputs via a mapping layer [250]. Variations on this model introduce various feedback paths (see for example Section 3.1.2), different sensing strategies and controllers [229, 364], or different mappings of gesture to sound output [148]. These models, while providing an accurate description of the technical implementation steps and of the flow of information in the system, do not account for the variety of movements that can result in identical input, and only consider 'active' sensing elements in an instrument's design, without taking into account a user's experience of the instrument which falls outside the remit of these models. Levitin, McAdams, and Adams [201] provide an analysis of control parameters in new musical instruments and consider how the features of a sound itself might be most appropriately mapped onto musical controllers. Their approach aims to provide a consistent language for mappings that can help inspire the design of new controllers that are well matched to the characteristics of electronic sound, moving away from the traditional keyboard.

The representation of an instrument through its control capacity has occupied much musicological research, as represented by the work of Baily [15] that was covered in Section 3.2.3.1. An interesting example of this comes from Hood [142] who created a series of 'organograms' that each illustrate the key control elements of an instrument including the number of tones available, the hand position and number of fingers to be used, the angle the instrument is to be held at and even whether the musician should be standing or seated (see Figure 3.7). This kind of musicological description of an instrument is key to the way that different instrument categories are understood and to how these instruments are arranged in organological history.

The problems that electronic and digital musical instruments pose to organology has also previously been explored in the work of Kvifte and Jensen [188] who propose that musical instruments should be understood and compared based on playing technique. Their approach to instrument classification can be seen in Tables 3.1 and 3.2 where the control of the instrument is broken down into three parameters (pitch,

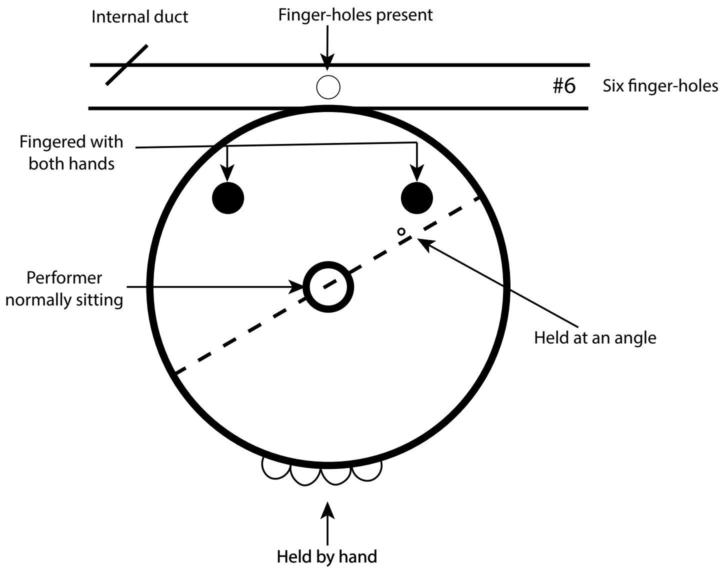



Figure 3.7: An adapted version of an organogram of a tin whistle from Hood [142] and quoted in Kvifte and Jensen [188]. This an example of describing an instrument through the musical control it makes available to a performer.

| Bagpipe | Pitch | | Loudness | | Timbre | |
|---|---|---|---|---|---|---|
| | A | D | A | D | A | D |
| Drone on/off | | | | ✓ | | ✓ |
| Fingering | | ✓ | | | | |
| Fingering type | | | | | | ✓ |
| Pressure on air bag | ✓ | ✓ | ✓ | | | |

Table 3.1: Tabular representation of the control dimensions of a bagpipe adapted from Kvifte and Jensen [188].

loudness and timbre) which can each be described as controlled in an analog (continuous) or digital (discrete) manner. With this model a clear picture of the kind of gestural language that exists within the dominant musical practice of an instrument can be seen at a glance and easily compared across different instrument types.

The focus of the models described above has generally been on instrument-like DMIs. Birnbaum et al. [31] offer an extension of these models that better accounts for the diversity of projects that can be broadly described as musical devices, extending the remit to include sound and interactive installations, sequencing tools, and musical games and toys alongside more classically instrument-like DMIs. They propose seven dimensions as a basis onto which these musical devices can be mapped as can be seen in Figure 3.8 (a). The result is a series of visual representations where clear trends in the shapes that

| Clavichord | Pitch | | Loudness | | Timbre | |
|---|---|---|---|---|---|---|
| | A | D | A | D | A | D |
| After-touch pressure | ✓ | | | | | |
| Choice of key | | ✓ | | | | |
| Striking force | | | ✓ | | | |

Table 3.2: Tabular representation of the control dimensions of a clavichord adapted from Kvifte and Jensen [188].

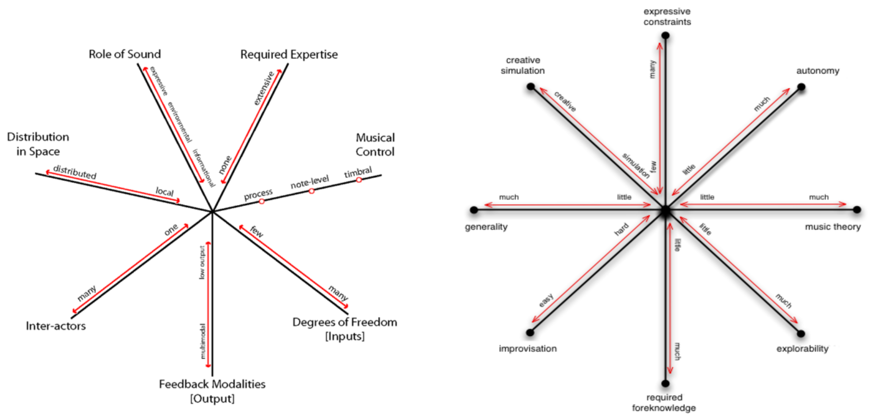
appear can be seen helping to classify and compare these diverse musical devices in terms of the following categories:

- *musical control*: whether the instrument's expression is at the timbral, note or score level.
- *degrees of freedom*: an axis that represents the number of input parameters that can be controlled.
- *feedback modalities*: the degree of real time feedback to the user.
- *interactors*: number of people involved in the performance of the device.
- *distribution in space*: the total physical area inhabited by the instrument.
- *role of sound*: informational, environmental or expressive.
- *required expertise*: representing the level of practice and familiarity needed in the performance.

Birnbaum et al. encourage authors who wish to employ this model to extend it and refine the parameters of the dimension space. The model in the form presented above remains focused on descriptive elements of the musical device: the space it occupies, how many people it is designed for, the degrees of freedom of control. The musical context that the instrument exists within falls outside of this model.

#### 3.5.1.2 *Definition through cultural significance*

Magnusson notes that the dimension space of Birnbaum et al. is largely occupied with phenomenological considerations [215] and, in a bid to better consider the cultural context of an instrument, builds upon this work by introducing the concept of *epistemic* musical instruments. This stance acknowledges how musical devices, interfaces and instruments are inscribed with knowledge of musical culture, that is the theoretical structures that users of such a device are forced to engage with [214], echoing material discussed in Section 3.4.4. Magnusson [215] identifies a lack of attention paid to the conceptual contexts that an instrument's design comes from, and in response offers a complimentary and overlapping approach that highlights the cultural aspects of an instrument's design. This is described as the *epistemic* dimension space (see Figure 3.8 (b)).




(a) Birnbaum et al.'s dimension space for musical devices [31].

(b) Magnusson's epistemic dimension space [215].

Figure 3.8: Dimension spaces for musical devices.

The value of Magnusson's system lies in its broad applicability: it can be applied to programming languages used for creating musical instruments as much as to the instruments themselves. In this research I am interested in the evaluation of a DMI from the perspective of the performer, without doubt the most important stakeholder in the process of designing and building a DMI [273]. We shall now consider the functionality of an instrument as it meets the performer.

### 3.5.2 *Playability*

The functionality of a DMI as it relates to the performer can be understood as its *usability* or *playability*. Wanderley and Orio [369] have provided a set of recommended musical tasks for the evaluation of the playability of DMIs. These include some fundamental musical elements such as arpeggios, scales and glissandi, the evaluation of which they recommend should be within a coherent musical context. They suggest that through the appropriate choice of musical task it is possible to evaluate features such as learnability, explorability, spatial controllability, and timing controllability. The maintenance of a valid musical context is an important aspect of evaluating DMIs and a way in which it is similar to experimental designs from music psychology (for example those on timing accuracy [94] or disruption [293] mentioned above) but differs from lab-based studies (on sensorimotor control [308] for example).

An evaluation of playability can take the form of measuring the movement of a performer and then using this data to compare performance under a number of different settings. Accuracy of tuning and timing are common points of analysis and follow more general methodologies from HCI on evaluation of task completion time or accuracy of target finding [273]. The quantitative data gathered for such an analysis can include sensor data from the instrument or motion capture data from the performer, that is then subject to later analysis deriving temporal accuracy, intonation, and gesture classification. Other approaches include the analysis of performance technique and its diversity between performers or with performers playing variations of an instrument [125, 389]. This type of analysis gives a wider picture of interaction with a DMI that usually removes the constraints of a musical task, and allows musicians to use an instrument for their preferred style of music-making.

DIVERSITY From the perspective of the instrument designer there are also a number of self-reflective assessments of a DMI's design that can be used to understand how it supports a performer's interaction. Jordà suggests that all musical instruments (DMIs being no exception) can be described in terms of their ability to allow for a diversity of musical styles and playing techniques [167]. In his framework there are three classes of diversity (Micro, Mid and Macro) and every instrument can be evaluated in terms of the level of diversity it has in each of these classes:

- *Micro diversity*: the degree to which an instrument can reflect very fine nuances in performance (related to control intimacy)
- *Mid diversity*: the degree to which two performances on the same instrument can differ
- *Macro diversity*: the capacity of an instrument to be used in a number of different styles and musical contexts

These classes allow an instrument's suitability to fit into a particular musical context to be assessed, alongside its capacity to support expressive performance.

LEARNABILITY In relation to the learnability of a DMI Jordà introduces the notion of *efficiency* which can be considered as the instrument's ability to transfer input gestures to musical sound [167]. Jordà posits that the kalimba, at least in the first stages of learning, is a more efficient instrument than the piano. Whereas a piano has many notes a kalimba has few and they are all the 'right notes'; its form intuitively encourages the performer to play with their thumbs; once the kalimba is held in both hands it is clear which thumb should be used to control which notes. So instrument efficiency depends on the relationship between the complexity of the input gestures and the complexity of the resultant sound output, and the ease with which a performer can get from one to the other. Jordà illustrates the balance between challenge, frustration and boredom by comparing these instruments, and suggests that new instruments should adhere to Wessel and Wright's 'low entry fee with no ceiling on virtuosity' [374].

A developed exposition of this mix of skill and challenge can be found in Csikszentmihalyi's concept of "flow": a psychological state that combines an increased sense of control with decreased self-consciousness [64]. The concept of flow describes a concentrated mental state in which a person is completely immersed in an activity: when a flow state is reached action and awareness are merged, and this often occurs in situations where challenge and expertise are balanced. Nash and Blackwell have applied theories of flow to the musical context, analysing interaction with non-real time software compositional tools [258, 259]. They observe that states of flow are reached when interaction is built on rapid feedback cycles in combination with mastery of the motor skills necessary for the musical task. The development of a flow state shares similar characteristics to the progression from explicit reliance on feedback in performance to a feedforward prediction state.

Pardue states "it is in learning where DMIs can have an inherent advantage over traditional instruments" [284, p. 45], proposing the term *complexity management* to describe the notion that instrument efficiency can be progressively managed over time in order to maintain a rewarding learning experience at all levels of expertise. By guiding the novice user towards less complex musical output, certain techniques can be isolated for technical practice. This approach could also provide more immediate access to less formal musical practice such as improvising or 'jamming' with other musicians, activities which McPherson and McCormick [241] show help with the level of cognitive engagement during musical practice. In DMI design, relating control strategies to those of acoustic instruments can be a method of leveraging existing expertise, although exactly which design elements are responsible for this remains unclear: as Cook points out "copying an instrument is dumb; leveraging expert technique is smart" [61, p. 1].

### 3.5.3 *Performer experience and instrument quality*

A number of techniques aimed at understanding performer experience have been applied in the evaluation of DMIs, and many relate directly to HCI techniques of evaluating user experience which can include questionnaires, comparisons, interviews, observations, interaction logging, and physiological measurements [44]. However some of these techniques can be applied to DMIs more easily than others: in HCI user experience is often assessed through the use of "think-aloud" methods, where the user talks through their thought processes as they are interacting with the system, reporting any difficulties in their interaction as they encounter them [359]. With a DMI this can break the flow of a musical interaction and disrupt a performer by distracting them from their control of the instrument. Strain, Shaikh, and Boardman [345] present an amended version of this technique that is better suited for realtime tasks that cannot be interrupted: participants are invited to reflect on their interaction after the event and are prompted by an experimenter to reflect on certain aspects of their interaction, or on certain problems that the experimenter identified they were encountering [345]. These structured interviews are then subject to further analysis in order to draw out themes.

In many cases a combination of quantitative and qualitative methods is used. Questionnaires that include quantitative input (continuous sliders, Likert scales) are commonly used to capture information about a performer's experience, often in relation to a comparative judgement. Choosing the right wording to have above a slider is a difficult task and often researchers use terms derived from quality measures used in the evaluation of acoustic instruments such as responsiveness, naturalness, richness [99, 317, 318] as we discussed in Section 2.3.1. With these kind of techniques there is a danger of conflicting interpretations of a specific descriptor between participants, and this is where the richer data set that can be gathered using qualitative techniques can help to provide clarification, specification and explanation for the ratings that performers give in a questionnaire. Qualitative data takes the form of recordings of structured interviews with participants and recordings of improvisations performed on the instrument. Thematic and lexical analysis tools can then be applied to these recordings such as discourse analysis [344] or thematic analysis [40]. In comparison to quantitative questionnaires, these approaches make room for subtle variations in participants' responses and for quantitative results to be extracted from unstructured dialogue.

**Reflective longitudinal evaluation** An increasing number of papers have begun looking at longer-term trends in the design of musical instruments, whether that be the life cycle of particular NIME instruments [254] or the experience of NIME performers over a longer period [253]. Similarly surveys of a certain field of musical practice are a popular technique for gaining an understanding of larger themes and trends in attitude, for example wide-reaching surveys on acoustic and digital musical instruments [218], overviews of sensor usage in DMI design at NIME over a ten year period [225], or analysis of the usage of the word *gesture* in the proceedings of NIME [162].

### 3.5.4 *DMIs as probes*

Evaluation and analysis techniques from design research that aim to encourage reflection on and participation in the design of an object can also be of great use to evaluating DMIs. The instruments created during this research where designed to act in a manner similar to 'technology probes' as proposed by Hutchinson et al. which are created in order to serve three goals: the social science goal of understanding the needs and desires of users in a real-world setting, the engineering goal of field-testing the technology, and the design goal of inspiring users and researchers to think about new technologies [150]. As such, each instrument was deliberately created to provoke reflection on a certain type of interaction or musical control: they were designed to encourage thinking about the dynamics of the interaction with the instrument. These reflections were captured through video interviews and 'think-aloud' demonstration sessions, while performance was captured by the sensors embedded in the instrument itself.

Technology probes build on Gaver et al.'s 'cultural probes' [104], and similarly they are designed to simultaneously elicit reflection on technology use and gather information about users. Cultural probes, however, are not necessarily technological objects themselves and rely heavily on external observation, whereas technology probes are designed to also gather a large set of rich information about the interaction that ensues [33]. Instead of being tool kits for gathering information about an everyday scenario, technology probes are low-fi technological devices that are designed to collect information around usage, explore issues of usability and most fundamentally to inspire users to reflect on the interaction and design [212]. Part of the appeal of the probe methodology in general is the flexibility with which it can be applied, the richness of the information it can gather, and the importance that is given to the design of the probe in terms of its provocative and disruptive power to a particular situation where design can intervene [33].

Technology probes have been used extensively in HCI over the last decade. Musical applications include the observation of compositional practices with novel interfaces [102], an evaluation of hand-controlled guitar effects [140], explorations of how an acoustic guitar can contain its own history through a digital archive that it builds around itself [22], a series of experiments around the D-Box hackable DMI [236] including how musicians appropriate its design [389] and how non-musicians come to terms with its design [237], and studies on the appropriation of highly-constrained musical instruments [125, 223]. Other wide-reaching approaches have included large-scale longitudinal interaction logging aimed at studying 'probe' music interfaces for composition in a real-world setting [258, 259].

In the work presented in this thesis maintaining a real-world setting was also an essential aspect of each study that took place, where the musicians were encouraged to improvise freely and perform musical tasks with the instrument in a setting akin to a recording studio or rehearsal room. The instruments themselves were all fully functioning with the main technological problems solved in their design, yet they were all simplified in a way, reduced instruments that allowed particular aspects of the interaction to be considered in isolation, and the musicians' attention to be focused towards a specific design factor.

### 3.5.5 *Designing patterns of movement*

Conceptualising how the movement of a performer and the mechanism of an instrument relate is a necessary step when designing new musical instruments. As discussed in Section 2.4.2, considering a musical instrument in terms of its affordance structure is a widely applied conceptual framing when designing musical interactions [59, 105, 216] and in design more generally [267]. Affordances, as outlined by Gibson, denote the possibilities for action that an object holds in a specific situation which exists relative to the action capabilities of an actor [113]. Gibson's affordances are contained in the physical object and do not change based on an agent's needs or skills. Although an object may invite certain types of action, whether the agent notices or ignores these affordances does not change them. They are not projected onto an object based on an agent's notion of what's possible, as with Norman's perceived affordances [267], but are held in an object. This distinction between objective and perceived affordances is key to the projection model that shall be presented in Section 3.6.

Performer-instrument interaction can equally be viewed through an instrument's constraints, that is the limitations of an instrument's design whether subjective, objective or cultural [216]. As opposed to affordances, constraints often are imperceptible at first and are discovered through exploration and experience. Every DMI can be considered as a structure of constraints that are decided by mappings and physical design, themselves considered a kind of compositional process, in the sense that the constraints of an instrument are necessary for a coherent expressive structure of musical possibilities to be created.

#### 3.5.5.1 *Expected, sensed, desired*

In Benford et al.'s *Expected, Sensed, Desired* framework [21] three categories of movement are presented from the perspective of the designer creating a sensing-based interface. *Expected* movements are all the body movements that could be considered as physically possible with a device. Some of these occur as a by-product of the sensing mechanism but are not directly sensed by it. Expected movements are pre-accounted for by the designer. *Sensed* movements are those that the device is able to capture and utilise through its sensing mechanism. *Desired* movements are the type of movements that a designer wants a user to perform in order to use the device, and can be considered as similar to the *pre-choreography* [204] of the device, as shall be discussed below.

Benford et al. propose that the most interesting lessons for designers lie in the boundaries between each of these categories and in the movements that fall into only one category [21]. They suggest that rather than aiming for full sensor coverage, designers should consider deliberately building 'rest spaces' into the experiences they create. In terms of electronic instruments, Bowers and Hellström [38] speak of an 'expressive latitude' in their designs. Space is deliberately left free for expressive body movements which are not sensed and have no technically-mediated musical outcome. Far from being 'dead zones', movements that are expected but not sensed can provide an important space of opportunities for adjusting, resting, preparing, and other performance-critical features of physical movement [21]. The projection model that follows seeks to underline the importance of these other performance-critical movements, beyond the input-output model [356], and to find a descriptive language for the design elements that underpin them.

#### 3.5.5.2 *Who controls who?*

When discussing the implications of embodiment for HCI and design research in Section 2.4.3 I introduced the notion of *embodied control*: how everyday action is *choreographed* during the use of technology. This theoretical framework re-evaluates the concept of control in the context of embodied interaction and human–technology choreographies. Tuuri, Parviainen, and Pirhonen [355] describe embodied control as consisting of three main sub-categories, or stances: *instrumental control*, *experiential control*, and *infrastructural control*.

INSTRUMENTAL CONTROL This relates most closely to the kind of models we have seen in Section 3.5.1. *Instrumental control* is the manner in which users' embodied interaction with a device provides them with control, and equally the manner in which the same device enforces control on the user [355]. This can also be described as the *pre-choreography* or choreographic inscription [204], which encourages movement configurations (both intentional and unintentional) through the device's design: desired patterns of movement are inscribed into the device. *Instrumental control*, then, relates to the choices made by the designer that enable or disable a performer to make musical sound with an instrument.

EXPERIENTIAL CONTROL While instrumental control refers to how the designer provides technical means of control to the user, *experiential control* switches viewpoint and considers the user's conceived control of the device, a device's *controllability*. This is about what users feel they can *do* with a device and relates to the material discussed in Section 3.5.2.

Tuuri, Parviainen, and Pirhonen [355] describe how users come to terms with their experiential control: it is formed through their interaction with the *push* and *pull* effects that an instrument displays. This is an interpretation of affordances that divides their perceived effect into those that either *repel* or *attract* the movements of a performer towards a goal. *Push effects* of a design are described as the way a user feels forced, guided or disabled in control. In the case of a DMI this could be how the instrument pushes performers to play in a certain way through conflict between action and desired result. *Pull effects* are feelings of being enabled in control and relate to an ease in conceiving how action relates to output. In Magnusson's discussion of constraints in DMIs [216] he suggests that the designer is often concerned with creating affordances, whereas the performer is more interested in navigating constraints: the design process may relate more to instrumental control even if experiential control is more operative to the person encountering that instrument.

INFRASTRUCTURAL CONTROL This concerns the way that a particular design can influence the infrastructures that partake in controlling human action, for example how an instrument speaks back to its musical culture and can intervene in and change it, echoing the material discussed in Section 3.4.4. The choreographies that form between a performer and instrument are not solely a result of how movement patterns are accommodated by an instrument's design: they are also shaped by the "physical and social infrastructure that facilitates, triggers, guides and orientates the dynamics of everyday movements" [355, p. 2]. The choreographies that emerge relate equally to how an instrument fits in, and has an effect on, the whole infrastructure that defines our encounters with our everyday environment [75].

## 3.6 THE PROJECTION MODEL

Earlier sections of this chapter have reviewed theoretical and analytical models of DMIs and have demonstrated the many ways in which an instrument can be represented and interpreted: as an information system through which data flows from input to output [148]; through its phenomenological relationship to a performer [31]; through its significance to wider cultural practices [215]. Over the course of this research project a conceptual model of performer-instrument interaction has been developed. This model shall be discussed in further detail in Chapter 7 where it is applied as a descriptive tool to a post-hoc analysis of the instrument used in Chapter 5. At this point in the thesis it is worth outlining the model as it will be employed in the discussion sections of each of the subsequent chapters.

In this model a performer-instrument coupling is considered as a projection from action to sound. This descriptive model helps to highlight how instrument design and performer action relate, and how idiomatic styles of playing are established on an instrument. The model brings together three elements of performer-instrument interaction:

- *Projection* relates to the objective possibilities of the physical coupling between performer and instrument
- *Aperture* is used as an objective descriptor of an instrument's design and the manner in which it transduces action to sound
- *Choreographies* relate to the patterns of action that form around an instrument

A performer projects their actions through an instrument into sound. In the process of projection the actions of the performer are refracted, brought into focus or blurred, and only a subset of their actions are able to pass into sound. Central to this model is the fact that in the process of transduction from action to sound 'information' about the performer's action, or in the analogy of the projection, light, is lost. Projection has been chosen as the primary metaphor of this model as it captures a sense of dimensionality reduction: in the passage from a higher dimensional space to a lower dimensional subspace information is necessarily lost and there is a minimal representation decided by the aperture.

#### 3.6.0.1 *Thinking from the instrument*

A physical characteristic of a musical instrument is that the gestures of a performer are in some ways transduced by the instrument into sound, as discussed in Section 3.2.3.1. Only certain gestures make musical sense with an instrument, and are selected for by the instrument's design. De Souza describes this as a performer's *body-sound coordination* forming through the correlations and invariance of sounding and kinaesthetic patterns: "both sound and action are facilitated and constrained by the instrument's affordances" [68, p. 15]. It has long been known that there exist gestures that are not directly sound-producing and yet are coupled to musical intention – see Section 2.2.1. Such movements provide indirect support to the sound-producing gestures (ancillary gestures [367]), and can relate to communication of musical or emotive concepts [163]. Many of these actions are neither sensed nor desired [21]: the instrument cannot capture them and the designer does not consider them as integral to musical performance.

As an example let us imagine it was possible to have a perfect motion capture system that could record every detail of a pianist's movement while playing. This recorded dataset would allow us to recreate the exact timing and velocity of key presses that the pianist performed, and from this it would be possible to reconstruct the whole performance in all its musical richness. If, on the other hand, we were presented with the same performance as captured by the keyboard – if we were given the exact timing of note onsets and the exact velocity of key presses – it would be impossible to recreate the body movement that led to the music. Although we might trace some information about the body movement from this second set of data (for example a certain note being played slightly later might indicate that the pianist has jumped from one area of the keyboard to another, or a quieter key press in certain positions might suggest that note fell under weaker fingers), it would be impossible to work back upwards from the instrument's reduced input to the body language that led to this input. A plethora of different 'choreographies' of bodily movement could have been used to create the patterns of note onset with velocity as sensed at the keyboard. This can be imagined as looking at a 2D rendering of a scene and trying to understand what combinations of 3D objects and light sources might have produced it.

There is an extensive discourse around body language in pianistic performance: fine control of the arms and shoulders, the exact degree to which one should raise the hands, the amount that the head sways with the phrasing of the music, the breathing patterns that accompany performance [285]. When it reaches the piano this rich vocabulary of body language is reduced to an action that is essentially the control of a discrete event plus a single dimension of variation: note-on plus velocity control [277]. Nonetheless, people believe in the value of, and spend many years developing, an elaborate and rich body language just to be able to project it down onto this interface, which is in many ways restrictive. The importance of this body language for the interpretation of musical performance is indisputable, and has been widely recognised [20, 74, 354] (see Section 2.2.1). When discussing the role of touch in artistic pianism, Doğantan-Dack provides a wealth of tactile descriptors from piano pedagogy that excellently represent the complexity of this space at the level of finger contact with key, such as 'sinking into', 'fusing with' the piano keys; 'clinging to the keys as to something soft, velvety or downy'; 'kneading the keys as if with silken fingers, as if moulding warm clay'; 'pressing the key as if grasping the hand of a friend with warmth' [74].

As far as the physical mechanism of the piano is concerned the only important part of a musical gesture is the part responsible for note onset: the attack, and this could be considered the meaning-bearing component of such a gesture from the perspective of the instrument. However the gestures that lead to the note onset (ancillary gestures [367]) evidently leave their imprint on the tone quality perceived by the audience and by the performer. According to Shove and Repp, listeners do not just hear the onsets of sounds because "attacks are nested events, constrained by, affected by, and thus lawfully specific to the performer's actions. To hear the attacks is to hear the performer move. The dynamic time course of these gestures is reflected in the resulting sound stream" [332, p. 60]. This echoes the material on the importance of bodily movement for the interpretation of music discussed in Section 2.2.1.

This disparity between the richness of potential action from the performer and the reduced representation of this action in the instrument is what led to the conceptual model of the projection. As well as through its technical function and behaviour (the perspective of the keyboard), an instrument can also be understood through the manner in which it encourages and even obliges a performer to move in certain ways: the *choreographies* of movement that it makes possible through its design [204]. Action as registered by the instrument is not the whole picture: it is important to recognise the complexity of the gestural input space when designing and this is where the projection model can assist.

### 3.6.1 *Projecting from action to sound*

The concept of *projection* aims to capture the transfer function from the broadest space of musical action to the musical sound that results from performer-instrument interaction. A projection is defined as a property of a performer-instrument coupling, and represents the objective space of possibilities between a body and a sound, grounded in the physical instrument. A representation of the projection model can be seen in Figure 3.9.

There is inherent 'information' loss in a projection: certain actions are selected as musically meaningful, given an instrument's design and the 'fit' of a performer's body to that instrument. Let us imagine projecting from a 3-dimensional space down to a 2-dimensional representation of that same space: certain movements in 3D are carried into the 2D representation, while others are shrunk, distorted or hidden altogether, some movements are brought into sharp focus while others are blurred. In this model the metaphor of projection serves as an analogy of how an instrument transduces action to sound, how

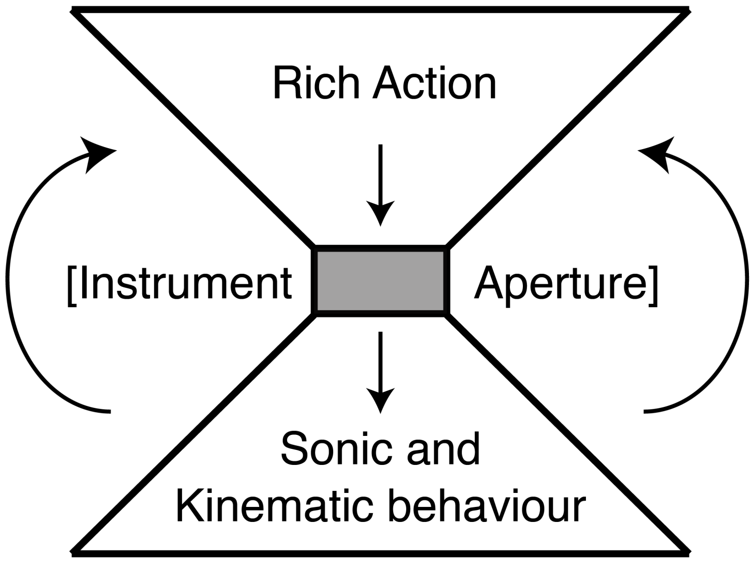



Figure 3.9: A representation of the projection model. The rich action of the performer is projected into the sonic and kinematic behaviour of the instrument via the aperture of the instrument.

traces of action are maintained in sound, and the inherent dimensionality reduction of this process.

A projection depends on the physical coupling of a performer's body and an instrument's mechanism, and so projections change based on factors such as body size and physical ability. Equally, different postures or ways of holding the instrument enable different projections, however in this model we limit the concept of the projection to the space of objective possibilities between the action of the performer and the constraints of the instrument, and so factors relating to training and style, although they have a strong influence on how a performer moves their body, are better discussed in the category of choreographies. Every Gibsonian affordance which produces a sonic result would be enabled by some aspect of the projection. It could be said that the projection of an instrument consists of the set of transfer functions for every available affordance: there is no projection that is not coupled to an affordance.

We can image, by way of an approximate analogy, that a lens is responsible for this projection. The characteristics of the lens decide on how action passes to sound: which actions are brought into focus, magnified, blurred or refracted. They also decide on the action space that is captured in the first instance through the lens's orientation. The rich action of the performer is focused and refracted and ultimately projected down through the *aperture* of the instrument. When discussing a particular instrument-player combination we describe the projection that results rather than attempting to describe the hypothetical geometric characteristics of the lens that is creating the projection.

### 3.6.2 *The aperture of an instrument*

An *aperture* in photography is a possibly adjustable mask that allows only a limited amount of light to pass, functioning in a manner similar to the pupil in the eye. In the projection model we can imagine that the action of the performer meets the aperture which controls the overall amount of action that passes, and can be considered as representing the point of minimal information flow in an instrument. The aperture has bottleneck-like characteristics in that it decides on the bandwidth of information passing through the instrument. Considering the point of minimal information flow in an instrument's design is imperative for understanding performer-instrument interaction and defining the aperture of an instrument should be recognised as an essential step in instrument design. The aperture is an intrinsic element of digital musical instruments and can be due to numerous different factors. When designing DMIs it is necessary to consider that the rich action of the performer will meet this aperture which decides on the quantity of information that is able to pass into sound.

### 3.6.3 *Developing choreographies*

*Choreographies* is a third term used to represent the patterns of action that form as a musician inhabits the projection on an instrument. This is a term taken from Tuuri et al.'s notion of technology-induced choreographies [204]. They propose choreographies as a means of describing how movements are used in the control of technology, and how reciprocally, movements are controlled by technology. Technology, from their perspective, contains pre-choreographic inscriptions, predefined patterns of action that are then enacted through interaction when the choreography is performed.

They suggest that embodied control can be best understood by switching our focus away from the technologies of an instrument to instead view control from the subjective and intentional viewpoint of a musician controlling an instrument [355]. It is via Tuuri's *push and pull effects*, as discussed in Section 3.5.5 that we move from projection to choreography: the subjective or experiential aspects of the performer-instrument relationship that lead towards certain actions and away from others. Choreographies are patterns of action that form around an instrument: a set of actions and gestures that the performer is likely to take. This is related to idiomaticity: performance practices which are well suited to the performer, instrument and musical context. Choreographies are observable in the external world, as shall be demonstrated in Chapter 7 but are the result of a subjective process and strongly culturally mediated: a violin and a fiddle have different choreographies even if the physical instrument may be identical.

Equally it would be possible to use the term *kinetography* to represent writing movement, as opposed to writing dance, and to distance ourselves from the colloquial usage of the term but we have chosen to remain with choreographies to align ourselves with Tuuri et al.'s definition.

#### 3.6.3.1 *Implications for evaluating instruments*

As with all frameworks the goal is not to be exhaustive and to cover every circumstance, nor to describe every kind of practice, but to illuminate aspects of how we design musical devices and how performers interact with them. The evaluation approaches and models presented in this section are not mutually exclusive and in fact should be understood as many different perspectives on the same landscape. The projection model is equally another perspective from which to view DMI design. The aim of this model is to draw attention to the choices made by designers, particularly design decisions relating to physicality and tangibility, and to highlight the projection from rich gesture to sound that is inherent to musical instruments.

The projection model seeks to find appropriate language to describe digital musical instruments and the way they encourage certain patterns of action. This model details the physical coupling between performer and instrument to highlight the role of tangible and haptic elements in the moment of contact between a body and an instrument. As argued by O'Modhrian [279], understanding this coupling is essential for digital musical instrument design, for which defining this coupling should be considered as a central tenet.

In the post-hoc analysis of the percussion instrument used Chapter 5 that is presented in Chapter 7 I shall illustrate an example of an application of this model to the analysis of a specific musical interaction context. I shall then discuss the advantages of this model when considering the influence of design variations on the choreographies that form around an instrument, particularly those related to physical aspects of an instrument's design that often fall outside the common sensor-mapping-sound model such as material characteristics, size and arrangement. In the discussion section of each of the subsequent chapters I shall also use the model to reflect on the findings from each study and the instrument design in each case.

## 3.7 CHAPTER SUMMARY

This chapter has discussed technology and design in relation to human haptic capabilities as explored in Chapter 2. We began with a discussion of sensorimotor skill and HCI's desire to capitalise on the sensorimotor capabilities of a user in the control of a computer. By tracing the historical waves of technology and theory in HCI, we identified that although digital technologies often come with increased power and competence in comparison to their analogue fore bearers, they often fail to utilise the sensorimotor capabilities of the user in control. This separation of control and action can result in the loss of rich physical feedback from the tool, feedback that is present in the case of direct operation. We moved on to review some of the technologies and applications, both musical and non-musical, that have been created to reintroduce this feedback. This work underpins the design study presented in Chapter 4 where an instrument with dynamic vibrotactile feedback was created.

In Section 3.2 we discussed the field of tangible and physical computing, a field whose goal is a rediscovery of the rich physical aesthetics of manual interaction. *Tangible* here takes on a wider meaning, one that not only refers to perceptual information available to the sense of touch, but also to this information's reception and meaning to the perceiving body (we can think of the difference between hearing and listening, one involves intentional derivation of meaning). Tangible computing puts theories of embodiment into practice in the design of physical interfaces. We reviewed some notable tangible interfaces and ended the section with a discussion of where an interface and an instrument differ.

*Control intimacy* emerges as a key concept for embodied interaction with musical instruments. Control intimacy can be understood as the *psychophysiological* distance between the input and output of a musical mapping: how the detail and nuance of a performer's movement is translated into the musical behaviour of an instrument. There is both a temporal and spatial component to control intimacy: the temporal regards the synchronisation and coupling between action and sound which is the main focus of the study presented in Chapter 5; the spatial regards the resolution with which a musician's movement is captured, the subject of the study presented in Chapter 6.

With Section 3.4 we took a step back from the mechanism of the instrument to discuss the physical supports that surround the sensing mechanism of an instrument. Through a further discussion of enaction and the embodied perspective we explored how physical aspects of an instrument's design can play a substantial role in its mapping, an important area of investigation for the work presented in Chapter 6 and Chapter 7.

This chapter closed with a discussion of evaluation techniques as applied to DMIs. Interpretations of some of these techniques are employed in each of the studies presented in Part ii. Here we weighed up the relative pros and cons of each evaluation technique and reviewed theoretical models of DMIs that aim to inform future design and act as a tool to evaluate existing instruments. The projection model of performer-instrument interaction was introduced, a descriptive tool for discussing physical aspects of design in relation to performer movement.

# Part II

# PRACTICAL INVESTIGATIONS

# 4

# INCORPORATING HAPTIC FEEDBACK

*This chapter is built on significant material from 'Navigation of pitch space on a digital musical instrument with dynamic tactile feedback' by Jack, Stockman, and McPherson, originally published in the proceedings of the Tenth International Conference on Tangible, Embedded, and Embodied Interaction, TEI 2016 [155].*

The experiment presented in this chapter investigates broad issues relating to *feedback* in DMIs, that is how an instrument responds to our action, and in turn how this response shapes our actions. The experiment revolves around a DMI with variable vibrotactile feedback strategies that relate to a performer's intonation within a continuous pitch spectrum. The instrument uses audio-rate actuation to create different patterns of vibration that relate to the actual audio output of the instrument and to the performer's actions. Its aim is to investigate how rich haptic feedback can be reintroduced to a DMI, questioning how auditory and tactile feedback relate during musical performance, and how dynamic vibrotactile feedback can support musical gestures and influence musicians' impressions of an instrument.

## 4.1 RELATED RESEARCH

In Chapter 2 we discussed the primary feedback channels of aural and haptic stimuli that are at play when performing with a DMI, and their cognitive bases. Haptic stimuli come in a variety of forms, and haptic perception – the umbrella term that encapsulates the sense of touch – covers two distinct categories: proprioception, which is the sensation of the movement and position of one's body parts, and cutaneous or tactile perception, which is related to the perception of stimulation of the cutaneous receptors in the skin [100]. Vibrations are one of the primary stimuli perceived by the cutaneous receptors under the skin's surface, and give us information about the shape and texture of an object as we move our hands across it, and also about the kinetic and acoustic energy involved in a musical action. Whenever the sense of touch is employed, the perceiver is in reality always using a combination of information from these sensory channels, in different ratios and weightings depending on the characteristics of the stimulus [342]. One can imagine playing a violin to understand how intertwined these two modalities are when performing with an instrument: the violinist may have a proprioceptive sense of their bowing arm moving the bow back and forth while this motion is continually monitored through the vibrations transmitted through the bow and felt in the bowing hand, as well as the vibrations felt through the chin and shoulder that hold the body of the violin. The study detailed below concentrates on creating vibrations that target the cutaneous receptors in the skin in order to guide the proprioceptive movement of the performer.

The experiment in this chapter particularly explores audio-tactile multimodal perception. Tactile and auditory perception behave in a number of similar ways: both respond to vibrational signals and can perceive amplitude, frequency and waveform, within perceptual ranges and JNDs, albeit with different ranges and sensitivities. In Section 2.3.4 there is a detailed review of how musical parameters translate to the sense of touch. The skin can be considered an extremely short-range, lo-fi, high-area ear covering the whole body, with its most sensitive parts in the hands and lips of a person. It is no coincidence that these are the parts of the body most commonly used to control musical instruments. The similarities in the mechanisms of these two sensory modalities mean that at times the same energetic stimuli are perceived by both senses: in the case of performing on a double bass, for example, the performer will be mostly perceiving the exact same mechanical vibrational energy through their ears, as sound, as they do through their body as tactile vibration.

In addition to being another route for perceiving vibrational stimuli, the tactile modality can also be used to receive other forms of information that are represented through patterns (both spatial and temporal) and that stand as symbolic representations for instructions of some type. Braille is the most widely known example of this that utilises static haptic feedback. With Braille, different patterns and arrangements of raised dots come to represent letters of the alphabet, numbers and punctuation marks. Similar material related to tactile maps was discussed in Section 3.4.2. In a similar manner, patterns of vibrations can also take on symbolic meaning and communicate information about direction, distance, and intensity, and carry instructions about actions to be performed.

## 4.2 RESEARCH QUESTIONS AND SCOPE

This study addresses Research Question 1:

*How can active tactile feedback be incorporated into a DMI and what influence does its reintroduction have on performer experience?*

I address this question by conducting a comparative experiment with a DMI with different programmable vibrotactile feedback patterns, aimed at assisting with intonation on a single-voice instrument with

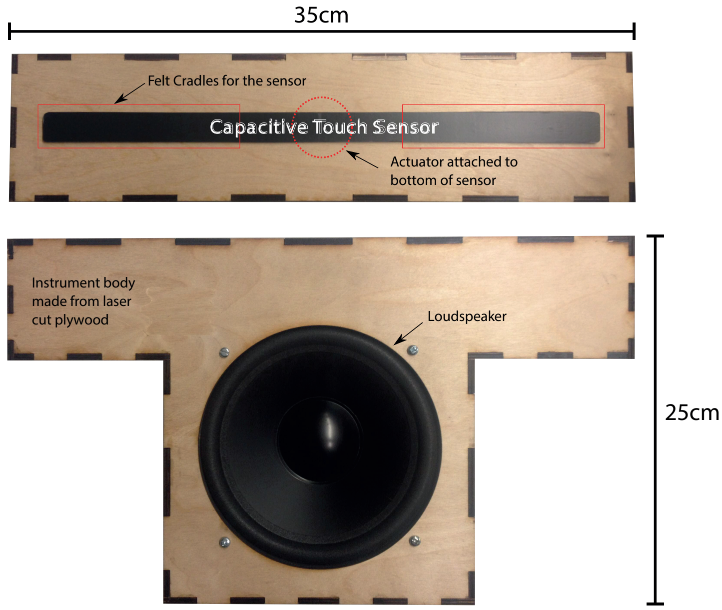



Figure 4.1: The instrument's top surface and front face with placement of actuator and felt cradles highlighted.

continuous control of pitch. In particular I address the following sub-questions of RQ1:

(a) What influence do different dynamic conditions of vibrotactile feedback have on musical performance?

(b) How accurately can musicians perform musical tasks on a DMI under each of these conditions?

(c) How does each of the vibrotactile conditions affect the musicians' impression of the instrument?

To answer these questions I assess the intonation of each musician while performing a series of musical tasks under each vibrotactile condition. This is combined with a gesture analysis of periods of free improvisation on the instrument with each of the conditions, and a thematically analysed structured interview. The analysis I perform combines techniques from HCI about efficiency of task completion and qualitative techniques that aim to assess the performer's impressions of the different feedback conditions.

## 4.3 Instrument Design

The technology probe (see Section 1.5) created for this study was a self-contained DMI. The pitch control of the instrument was constrained to a continuous range of two octaves. The instrument was able to vibrate in different ways depending on how it was played. The instrument was a battery-powered stand-alone embedded device created using the Bela platform (as discussed in Section 1.5). Bela

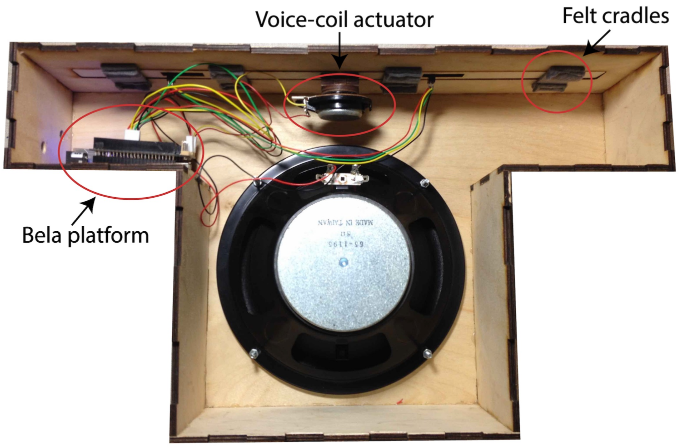



Figure 4.2: The instrument's inners, showing the actuator attached to the bottom of the touch sensor, the Bela board, and the back of the speaker.

deals with the touch sensing and the generation of audio and tactile signals, and also provides amplifiers which are ideal for driving a speaker and actuator.

The instrument itself (Figure 4.1) is T-shaped, and consists of a wooden body that contains a 20cm speaker for audio output, a 40cm-long capacitive touch surface for measuring finger position, a vibration transducer mechanically coupled to the touch sensor for haptic feedback, and the Bela platform for audio and sensor processing and data logging. The instrument is designed to be played with the performer sitting down and the instrument resting on their lap with the speaker hanging between their legs.

### 4.3.1 *Sensors and audio mapping*

The form of the instrument is inspired by the D-Box [236], but with a reduced control space: the performer is only able to control the frequency and amplitude of a monophonic synthesiser. The control surface of the instrument consists of two 20×2.5cm capacitive touch sensors derived from TouchKeys [238] on the top plane, that sense touch position and contact area (which roughly corresponds to finger pressure). The two sensors are joined into a single continuous strip that is 40cm long and is coated in a thin layer of vinyl that helps interpolate and smooth position readings from the multiple capacitive pads on each sensor.

The synth voice is a sawtooth wave with a frequency following bandpass filter ($Q = 0.5$). A simple one-to-one mapping relates touch location to frequency and contact area to amplitude. Frequency is mapped logarithmically, with a range of two octaves plus one semitone centred an octave below middle C (C3). The centre point can be felt as a small ridge where the two touch sensors join. The distance between semitones is approximately 1.6cm. Audio is sampled at 44.1kHz, and touch data is collected at 200Hz. Audio synthesis and vibrotactile control algorithms were built in Pure Data (PD)[1] [300] which was converted to C++ code using the Heavy compiler[2].

### 4.3.2 *Vibrotactile Feedback*

Vibrotactile feedback in a DMI is often provided by an eccentric mass motor, where frequency of rotation and intensity are coupled together. The instrument created for this study uses a voice coil actuator to provide audio-rate vibrotactile feedback. The relative merits and limiting factors of different strategies for providing vibrotactile feedback were discussed in Section 3.1.3. I was interested in creating a source of haptic feedback that could be treated as critically related to the audio output, providing information that is both time and space dependent. The feedback routines I designed into the instrument can provide information about zone borders and discrete regions of the instrument, and do so in a way that relates to the performer's continuous movement on the sensor and to the musical behaviour of the instrument.

actuation Tactile feedback was produced using a HiWave HIAX 25C10-8/HS actuator that was firmly affixed to the bottom of the touch sensor strip as can be seen in Figure 4.2, with the magnet able to vibrate freely inside the box. The sensor strip, as a result of being coupled to the actuator and being suspended on a felt cradle attached to the top panel of the box, was also able to vibrate freely. Although the primary goal of the vibrotactile feedback in this instrument was to target the performers' fingers as they played on the touch sensor, vibrations created be the actuators could be readily felt in the whole body of the instrument.

The actuator is driven with an audio-frequency signal, similar to a speaker. The frequency of tactile actuation ranged across two octaves from approximately 65Hz to 260Hz, following the frequency of the main audio output. The range of frequencies were chosen to specifically target the Pacinian receptors whose range of sensitivity is between approximately 50Hz and 350Hz [100]. The Pacinian receptors are fast to adapt and respond to vibrations. The nerve endings of the Pacinian corpuscles have been shown to fire in a way that is similar to the behaviour of the nerve endings in the basilar membrane of the

1 https://puredata.info/
2 https://enzienaudio.com/

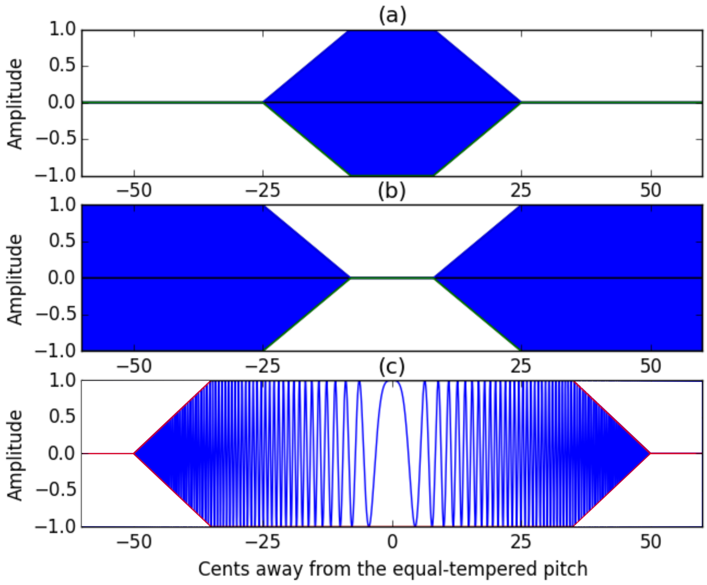



Figure 4.3: (a) amplitude envelope for the *vibrations when in tune* condition (b) amplitude envelope for the *vibrations when out of tune* condition (c) representation of beat frequency modulation with amplitude envelope in red for the *beat frequency* condition, mathematical description in Equation 4.2 and Equation 4.3

auditory system: they are believed to process stimuli by acting as multiple band-pass filters [220]. As with hearing, touch does not respond equally to all frequencies: just as we have equal-loudness curves for hearing [81], there are equal-intensity curves for vibrotactile perception [185] which display similar variation across the frequency range. To ensure equal intensity across the frequency range I applied approximate equalisation via a 4-band parametric equaliser which boosted the lower frequency bands and was a part of the software chain running on Bela.

conditions I considered four tactile feedback conditions that each generated a signal for the actuator based on touch location. The purpose of the feedback was to communicate to the performer when they were on or near a diatonic pitch of the C major scale. The conditions were:

1. *No feedback.* This was the control condition. No signal was sent to the actuator, but vibrations from the speaker can still be felt through the case and sensor, albeit at a substantially lower intensity than the other conditions.
2. *Vibrations when in tune.* In this condition, the vibration actuator engages when the pitch is near an (equal-tempered) diatonic note. As seen in Figure 4.3 (a), the amplitude of the vibrations is 0 when the note is more than 25 cents out of tune, ramping up to maximum intensity when the performer is within 8 cents

of the correct pitch. An in-tune area of 8 cents was chosen as a compromise between sensor size and frequency range. As we had 1.6cm for each semitone the in-tune area spanned 0.26cm. On reflection a smaller range would have allowed for more accurate feedback to be tested as shall be discussed in 4.6.

The frequency of the actuator signal is that of the correct pitch (i.e. the closest equal-tempered note), which is close but generally not identical to the frequency played through the speaker. The waveform was a sine wave. Compared to other work with fixed-frequency actuation signals [387], I chose to match the frequency to the nearest note to reinforce the relationship between tactile and audio output. This also led to occasional gentle beat frequencies due to the difference in frequency of actuator and speaker output. The beating effect inspired the deliberately exaggerated beats of Condition 4.

3. *Vibrations when out of tune.* As shown in Figure 4.3 (b), this feedback condition applies a similar technique to Condition 2, but reversed: the actuator is at maximum amplitude when the performer plays a note more than 25 cents from the nearest diatonic pitch, reaching zero amplitude when the pitch is within 8 cents of the target. The frequency of the actuator signal is identical to the frequency played through the speaker and the waveform a sine wave.
4. *Beat frequency.* This condition uses the difference between the target note and the played note to create haptic 'beating' [385], as a feedback strategy that emerges to the perceiver through its dynamic behaviour. This takes inspiration from the accounts of double-bassists mentioned above [96]. The beats are generated by interference patterns between two closely-tuned oscillators, according to the following trigonometric identity:

$$\sin(u) + \sin(v) = 2\sin\left(\frac{u+v}{2}\right)\cos\left(\frac{u-v}{2}\right) \tag{4.1}$$

Because the audio frequency differences involved are small, especially in the instrument's lower octave, the natural beat frequencies will be quite slow. To exaggerate the effect, I applied a warping to the reference frequency as follows:

Let $f_{spk}$ be the frequency of the speaker output, $f_{tuned}$ be the frequency of the closest diatonic pitch and $f_{beat}$ be the desired beat frequency at 50 cents away from the tuned pitch. Then $f_{ref}$ is the reference frequency used in the calculation of the warped beat frequency

$f_{warpedbeat}$. The $\pm$ in Equation 4.3 depends on whether $f_{speak}$ is above or below $f_{tuned}$. The calculation is as follows:

$$f_{ref} = (2f_{beat} - 1)|f_{spk} - f_{tuned}| \tag{4.2}$$

$$f_{warpedbeat} = (f_{spk} \pm f_{ref}) - f_{tuned} + 2 \tag{4.3}$$

I chose to create a beat frequency of 60Hz when the performer was 50 cents away from the tuned note, reducing to beating of 2Hz when the performer played perfectly in tune. The amplitude envelope applied to the actuator signal is shown in Figure 4.3 (c). Though on acoustic instruments the beating disappears entirely when two notes are precisely in tune, to implement this behaviour would create a subtle problem: depending on the relative phase of the two signals, the in-tune condition could be either a maximum or a minimum in amplitude, which could be confusing. Instead I chose to limit the minimum beat frequency to 2Hz so that performers would be able to feel a slow pulsing when in tune.

Referring back to Equation 4.1, we see the beating is a form of amplitude modulation. Here the frequency of the modulator is the beat frequency, the primary frequency that the performer perceives, which ranges from 2Hz to 60Hz. The frequency of the carrier is also variable (the mean of the reference and speaker frequencies) and provides additional haptic information. Since the beat frequency alone does not distinguish whether a note is sharp or flat, I hoped this additional carrier frequency information would help the performer distinguish between these cases.

## 4.4 STUDY DESIGN

### 4.4.1 *Method*

My experimental method was derived from Berdahl et al.'s study on pitch selection with force-feedback haptic assistance [26] which itself builds on the work of O'Modhrain [272]. Whereas these previous studies used haptic force-feedback to push the performer's hand onto the right pitch, this instrument requires an active correction from the performer in reaction to the vibrational patterns they perceive through their fingertips and thighs. To assess the impact of tactile feedback on pitch accuracy, I asked participants to perform a series of pitch selection tasks and play melodies using each tactile feedback condition. Alongside the musical tasks participants were also given

Figure 4.4: Participant performing on the instrument. The instrument is hidden from sight by a wooden panel.


Figure 4.5: The second simple melody that participants had to perform. Transposed up one octave for ease of reading.

periods of free improvisation on the instrument under each condition. Their performances were video and audio recorded, and they were also interviewed afterwards.

### 4.4.2 *Participants*

I recruited 10 participants (7 male / 3 female) aged between 26 and 45 years from my university department. A requirement for taking part in the study was that participants were competent musicians. All ten identified themselves as musicians; eight participants had 10+ years of instrumental experience, and the other two participants had 3+ years each. The most common primary instrument in the group was the guitar (5 participants) followed by piano (3 participants) and violin (2 participants). 8 of the participants had experience with electronic and/or digital musical instruments. The study lasted around one hour and fifteen minutes per participant and met ethics standards according to my university's ethics board.

### 4.4.3 *Experimental set-up*

The study was conducted in a sound-isolated recording studio at Queen Mary University. The experimental set-up consisted of the musical instrument and a wooden panel which hid the instrument from sight while the participant completed the task, as can be seen in Figure 4.4. The instrument produced sound through the speaker, but the participants wore noise-cancelling headphones during the experiment, through which they could hear the same audio signal played by the speaker. After a pilot study it was decided that the headphones were necessary to avoid any residual sounds from the tactile actuator influencing their performance. The speaker was left on at a lower amplitude for all conditions as the vibrations that it produced were judged as imperceptible on the touch sensor.

### 4.4.4 *Procedure*

The study procedure was as follows:

1. The study began with an introduction to the instrument without tactile feedback. The procedure of the experiment was explained, and they were given loose guidelines about how to play the instrument: only one finger was to be used on the sensor, although they were free to use any finger and change fingers and hands as they saw fit. Participants were also encouraged to rest their other hand on the body of the instrument to enhance the perception of tactile feedback.
2. Participants were then given ten minutes of free improvisation on the instrument to familiarise themselves with its range and behaviour.
3. After familiarisation, every participant performed with all four feedback conditions, presented in a counterbalanced random order between participants to minimise learning effects. The scales, arpeggios and melodies I will detail below were however presented in the same order.
4. For each feedback condition, the participant was first given a ten minute period of free improvisation followed by a note-finding exercise. Six single tones were played in isolation, and the participant had to match the pitch on the instrument. This task served two functions: to improve the participant's familiarity with the instrument under a given feedback condition, and to provide me with a metric of their pitch-finding skill.
5. The participant was then asked to play a series of scales, arpeggios and melodies. The scale was a 2-octave C major scale, first ascending and then descending. They could rehearse this until

comfortable and it was then recorded. A similar procedure was followed for a 2-octave C major arpeggio.

6. For the melodies, they would first listen to a recording and be provided with a score (e.g. Figure 4.5). Participants were advised to pick an appropriate tempo that allowed them to maintain accurate tuning with clear stable notes as well as a steady beat. After practising until they felt comfortable, they recorded the excerpt three times and chose their preferred take.
7. Upon completing the above tasks for a feedback condition, the participant completed a questionnaire about that particular condition, which included questions about their perceived tuning accuracy, their comprehension of the feedback and the mental effort it required. At the end of the study, the participants filled out an exit survey asking, amongst other questions, what their favourite feedback condition was. A short structured interview concluded the study in which participants were simply asked to justify their preference of feedback condition.

### 4.4.5 *Data collection*

Alongside the questionnaire results I also collected performance data from the instrument: the speaker signal, the actuator signal, touch position on the sensor, contact area on the sensor, the computed closest equal tempered note and the computed beat frequency for the beat frequency feedback condition. The two audio streams were sampled at 44.1kHz while the sensor streams were normalised to a range of $[-1, 1]$ and upsampled from 200Hz to a rate of 22.05kHz so we could treat the sensor streams as audio for later analysis. Audio and video were recorded throughout.

## 4.5 FINDINGS

This section presents the quantitative measures of performance under each of the feedback conditions, followed by a review of the reasoning that the participants gave for preferring or disliking each of the feedback conditions.

### 4.5.1 *Analysis method*

My analysis methods combined both quantitative and qualitative techniques. A primary interest was the quantitative measures of accuracy of performance, in this case intonation judged via mean absolute pitch error. I also assessed the temporal impact of each condition on performance by comparing the time taken to complete each task. The influence of each feedback condition on performer experience was also of interest and so I analysed their responses to the questionnaires alongside the structured interview conducted with each participant. Finally an analysis of emergent gestures in relation to each feedback condition is presented. This analysis used a combination of observation during the experiment and analysis of the sensor feeds recorded on the instrument.

### 4.5.2 *Mean absolute pitch error*

Based on the position readings from the touch sensor I compared the pitch of the actual note played by the performer against the target equal-tempered pitch, which was measured in semitones (logarithmic with respect to frequency). To begin, the performances were manually annotated on a DAW, first segmenting into notes, then identifying the region within each note when the performer settled on a stable pitch. In order to do so each of the sensor recordings were imported into a DAW as .wav files and visually analysed. A mean pitch value was then calculated for each stable section by averaging the pitch across the stable region with a python script. The absolute difference between this value and the target pitch values for each melody was calculated. This gave a measure of mean absolute pitch error for every note performed under each tactile feedback condition.

I used accuracy on the single-note pitch matching test as a screen for the reliability of the rest of a participant's data. Figure 4.6 shows that there was at times large differences between participants' performance. I excluded participants who achieved less than 80% accuracy in the pitch-matching task. One participant was excluded (Participant 7): as can be seen in Figure 4.6 their performance across most of the conditions was markedly worse than other participants.

Figure 4.7 (a) shows the medians and IQRs of the absolute pitch error for all participants while performing both melodies. I performed a paired t-test to assess the significance of the difference in absolute pitch error. I found a significant difference between 'beat frequency' and 'audio only' ($F(2,240) = 5.1, p < 0.05$) and a marginally significant difference between 'vibration when in tune' and 'audio only' ($F(2,240) = 4.5, p < 0.08$). This suggests that the different feedback conditions influenced pitch selection accuracy, although this is tempered by our relatively small sample size.

### 4.5.3 *Timing impact*

The average difference in timing for all participants and melodies was calculated by measuring the duration of each melodic passage minus

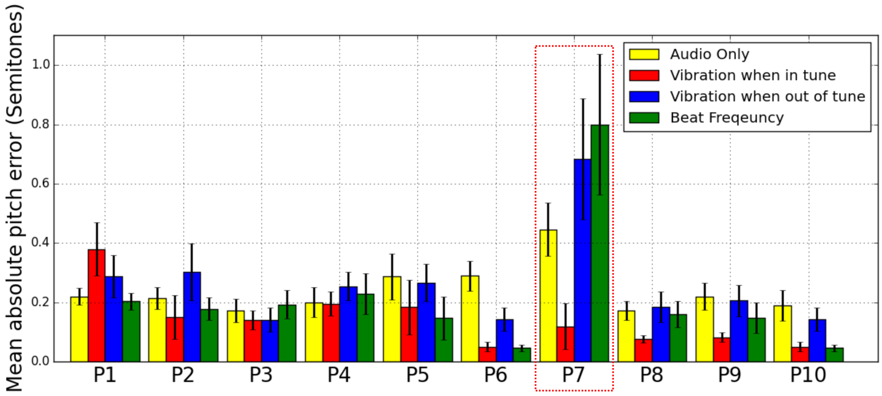



Figure 4.6: Mean absolute pitch error for each feedback condition for each participant for melody 1. Error bars represent standard error. The excluded participant is highlighted.

the first and last note. Figure 4.7 (b) shows that the 'vibrations when in tune' and 'vibrations when out of tune' were on average performed faster by participants than the 'audio only' condition (83% and 73% of the duration of the audio only condition). The 'beat frequency' condition however was generally slower than the 'audio only' condition (111% of the tempo), however no significant difference in timing impact for feedback condition across participants was found.

To examine temporal performance in more detail I then measured the duration of stable notes and the duration of the gestures used to reach them, the latter of which can be seen in Figure 4.7 (c). In the 'beat frequency' condition there is a generally smaller mean stable note duration than in any of the other conditions even though the time taken to perform each of the melodies is generally longer: the *searching* gestures before a stable note appeared to be the longest for this condition, yet again no significant difference was found across participants.

An additional factor that I took into account when analysing these results is the musical ability of the participants. I noted that participants who performed well in the initial note finding tasks went on to perform with a higher degree of pitch accuracy under all feedback conditions. As can be seen from the difference in performance between participant 4 and 7 in Figure 4.6, performance varied significantly from participant to participant. I took this difference in performance into account by excluding the performance data from participants who achieved less than 80% accuracy in the initial note finding task (one participant).

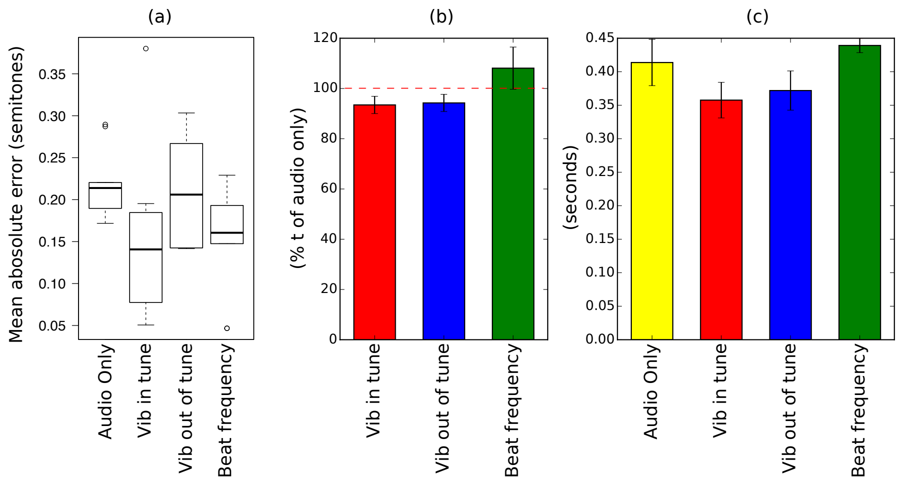



Figure 4.7: (a) Mean absolute pitch error for each feedback condition across 9 out of the 10 participants for both melodies. (b) Percentage of mean tempo in relation to the audio only condition. (c) The mean searching gesture duration (time before reaching a stable note). Error bars represent standard error.

#### 4.5.3.1 *Length of note-searching gesture*

A second measure I took was that of time taken to reach the stable note under each feedback condition, (length of the note-searching gesture). After witnessing the participants' performances I hypothesised that the length of time spent searching for the correct note would be increased with the tactile feedback as the amount of sliding in order to select the correct pitch seemed increased.

In fact my results show the opposite, as can be seen in Figure 4.7 (c): the length of time searching for a note was on average marginally reduced for two of the three feedback conditions in comparison to audio only feedback. It was only for the beat frequency condition that the searching time was on average longer than the audio only condition. For both 'vibrate when in tune' and 'vibrate when out of tune' the length of time spent searching for the tuned note was shorter than with the audio-only situation. When looking for trends across participants no significant difference was found for the length of each note-searching gesture suggesting that the observed trends may not hold in every case.

#### 4.5.3.2 *Learning effects*

The effect of learning was examined by comparing the pitch accuracy from the first and last feedback conditions each participant encountered. For 5 of 9 participants, the feedback condition that they performed with *first* was their most accurate, suggesting that learning effects due to the order of conditions are minimal in this study. The counterbalanced random order of conditions and the opportunity for practice on each melody also help reduce bias from learning effects.

### 4.5.4 *Playing Technique Observations*

Aside from the instruction to play with one finger at a time, the participants were not instructed to use a particular playing technique. I observed a wide variety of playing techniques that varied both among participants and feedback conditions. Techniques included the following: sliding from note to note on a single finger; various fingering positions on the sensor using different fingers with sliding used for small corrections or jumps of a large interval; detached playing with one finger where the performer lifts the finger between each note.

These examples show how each feedback condition engendered its own playing techniques that participants discovered through experimentation with the instrument. This finding is similar to that of Marshall and Wanderley [226, 229] during an equally open-ended study which evaluated sensor choice for parameter modulation in DMIs In Marshall et al.'s evaluation of the suitability of different sensor types for the control of different parameter modulations, they find that just as different instruments may naturally elicit their own specific set of movements and gestures, each of the sensors tested naturally elicited certain movements [229]. Furthermore, when suggesting how to derive a measure of sensor suitability for a specific task they highlight the importance of taking into account both the subjective measurements of user preference and the more objective measurements of accuracy and stability in performance.

#### 4.5.4.1 *Characterisation of note-searching gesture*

For each participant and feedback condition I noted a general trend towards certain styles of playing technique that were used to search out the correct note. The following observations aim to characterise some general trends in the sensor data that was captured under each feedback condition and are derived from multiple participants.

AUDIO ONLY The note-searching gesture for the audio only feedback setting can be characterised by a vibrato-like motion that overshoots and undershoots the note, oscillating with decreasing depth until settling on the stable correct pitch (see Figure 4.8). This technique is generally equatable to similar note-searching gestures used in string instrument playing, singing, or other instruments with continuous pitch control such as the trombone, or theremin [45]. Maintaining correct intonation on instruments with continuous control of pitch is a difficult task and usually requires the comparison of the played note to a reference note on the instrument (for example an open string on a violin) [284]. In string instrument performance vibrato is essential to developing an expressive tone but can also used to mask imperfections or uncertainties in tuning [109]. The frequency of observed vibrato was generally high (approx 4Hz) with a range of up to 0.2 semitones at times. As expected with this condition, the participants were relying on their ears for intonation.

vibration when in (and out of) tune With the vibrations when in tune condition the note-searching gesture generally contained an overshoot but was not followed by an undershoot like in the vibrato observed with audio only (see Figure 4.8). The participants wait for the vibrotactile feedback to be engaged, at times continuing beyond the feedback region but quickly returning to the point with feedback. With this condition there was a general trust in the feedback being provided which at times resulted in participant settling on a slightly out of tune note due to the range of the feedback (8 cents below and 8 cents above). This can be seen in Figure 4.8 in the second example for the vibrations when in tune setting. The aural feedback became less important once the confirmation was delivered. Similar behaviour was observed for the vibration when out of tune condition, but with the absence of feedback being the signal of confirmed intonation.

beat frequency With the beat frequency condition the whole note-searching gesture was slowed down (see Figure 4.8). Performers were much more concentrated on the feedback and this could grant them more accuracy but with the consequence of slowing down their performance. This could also be due to the unfamiliarity and complexity of this feedback pattern in comparison to the others.

#### 4.5.4.2 *Emergent playing techniques*

An additional performance technique that was observed for the 'vibrations when out of tune' feedback conditions was a method for searching out the correct note before playing (Figure 4.9). In this case the performer would lightly touch the sensor and adjust their position in response to the tactile feedback before applying more pressure and increasing the amplitude of the audio output to its full level. It is worth noting that the finger movement on the sensor in Figure 4.9 happens over a distance of ~3mm and lasts ~200ms. I did not mention the amplitude mapping to the participants – again this technique appeared as an intuitive response to the feedback condition.

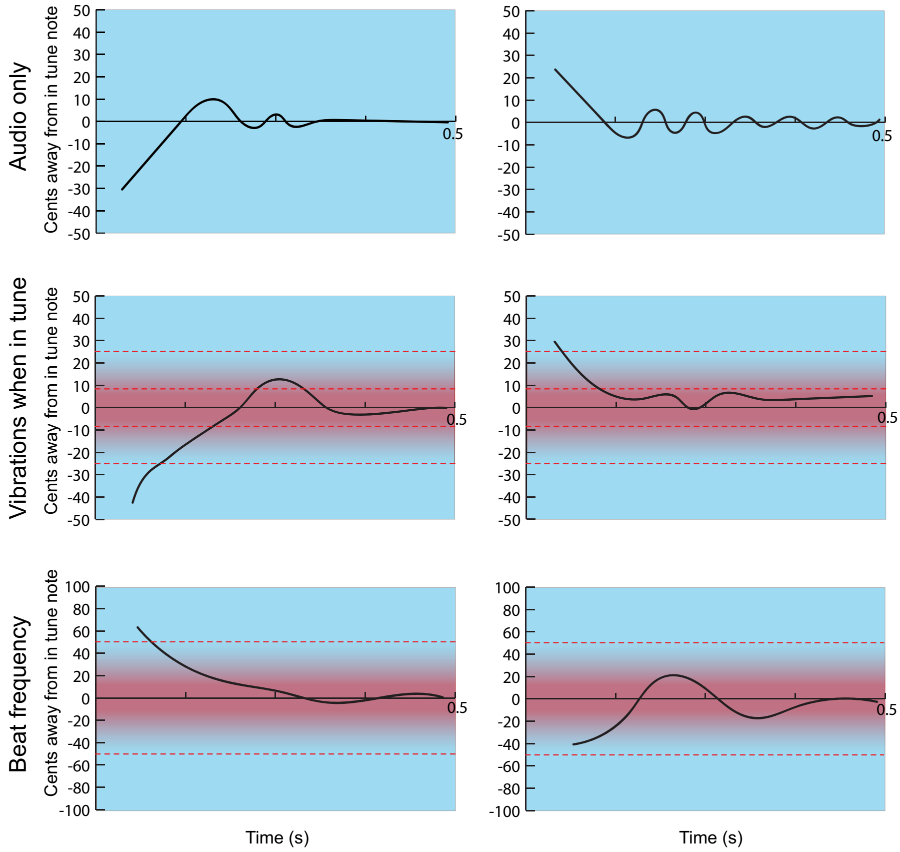



Figure 4.8: Some examples of the note-searching playing techniques that were used by participants in relation to the feedback settings. Dashed lines represent the point at which feedback started to be engaged and the point at which it was at its highest.

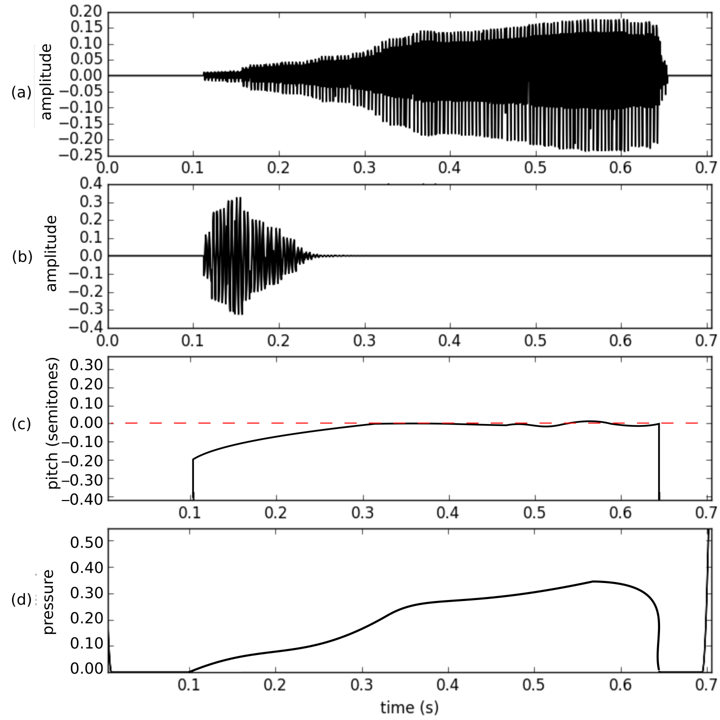



Figure 4.9: Note-searching playing technique, shown here for the 'vibration when out of tune' feedback condition. Shows the relationship of finger position and pressure to tactile and audio feedback. (a) audio channel, (b) actuation channel, (c) pitch from the sensor, (d) pressure (touch size from the sensor).

### 4.5.5 *Participant tactile feedback preferences*

A summary of the results from the surveys is presented in Table 4.1. In general, the 'vibrations when in tune' condition scored best on most metrics, though all tactile feedback conditions appear to reduce participants' reported ability to maintain their desired tempo.

Reports of preference and impact of the different feedback conditions varied greatly between participants. Table 4.1 shows the responses to each question on the survey for each participant and does not demonstrate a clear pattern in the rationale of the group as a whole. In terms of general preference there was no clear favourite feedback condition across all participants, as can be seen in the bottom row of Table 4.1. However for 7 of the 9 participants, the preferred feedback condition also yielded their most accurate performance in terms of intonation. Participants were asked to explain the reason for their preference and there was also great variety in the logic behind choosing a favourite.

vibrations when in tune For participants who preferred 'vibrations when in tune', they stated that they liked the affirmative nature of the feedback, that they were provided with a clear signal of when they were playing in tune and could play detached notes and know immediately whether they were playing the correct note. Here is a selection of quotes from the structured interviews:

> *Larger intervals were a bit easier as you knew there was a target to aim for (the vibrations). I also slid a lot and tried to count the notches to guess intervals which was almost possible though a bit tiring, better to estimate the note and see if it's there.* P10

> *It didn't get in the way of my playing at all. I was also able to press very lightly and check whether I was in tune then push harder when I was sure, this seemed quite natural to me.* P9

> *I became lazier and relied too much on it, more than on my ear* P6

> *In order to find notes you would have to rest for a bit longer and move your finger in both directions. So I scrapped it basically and tried to play the notes according to perceived distances. But tactile feedback helped me to find certain notes when I needed them.* P4

> *I found a portion of my concentration was taken in just looking for the feedback and forgot to concentrate on the music at times!* P1

vibrations when out of tune Participants who preferred 'vibrations when out of tune' stated that they preferred being 'buzzed' when out of tune as this reflected the way they would normally think

| Survey Question | Feedback Condition | | | |
|---|---|---|---|---|
| | *Audio* | *In Tune* | *Out of Tune* | *Beating* |
| How successfully did you play in tune? (1: Very badly – 10: Perfectly) | 4.2 (2.3) | **5.82** (1.2) | 4.78 (3.0) | 5.1 (2.8) |
| How hard was it to play in tune? (1: Very easy – 10: Very Hard) | 6.2 (3.2) | **5.2** (1.5) | 6.8 (2.3) | 6.0 (0.8) |
| Were you able to maintain your desired tempo? (1: Not at all – 10: Completely) | **7.3** (1.2) | 6.1 (1.3) | 5.0 (2.8) | 4.8 (1.9) |
| How mentally demanding was the tactile feedback? (1: Not at all – 10: Very) | N/A | **5.9** (1.2) | 6.44 (3.2) | 6.7 (2.1) |
| How much did the tactile feedback assist tuning? (1: Not at all – 10: Very) | N/A | **6.9** (1.5) | 5.5 (2.1) | 6.5 (1.3) |
| Which was your preferred condition? (Number of participants) | 0 | 4 | 3 | 3 |

Table 4.1: Summary of the responses from the survey conducted at the end of each feedback condition. Mean and (standard deviation) shown. The low and high limits of the 10 point metric are listed under the questions in the left hand column.

about pitch selection on their instrument: when they are playing in an ensemble and are in tune they don't think about their tuning; it is rather when they are out of tune with the ensemble that they become aware of their tuning and know they must correct it.

> *I was able to gently try and play the note and depending on whether there was feedback or not I could adjust my finger position. If I felt no vibration I knew I was in tune. This worked best on the scale where I had a pretty good idea of the intervals / structure of what I was playing. It also worked on the melodies but to a lesser extent as sometimes I would get lost and not know where the next note was, maybe I was closer to the semitone below and would go towards there.* P10
>
> *I knew straight away when I'd got the note right (without sliding), so that helped 'calibrate' me for the next note.* P3
>
> *Somehow this made more sense to get feedback when making a mistake or small error.* P1
>
> *Because vibration meant out of tune I felt it stressed me more than not getting tactile feedback in the first case.* P8

beat frequency Reasons for selecting the 'beat frequency' tactile feedback as the preferred method had to do with the variety and amount of information it provided. Participants who preferred it stated that this condition had the most potential for long-lasting engagement, as it allowed micro adjustments to tuning to be performed and helped maintain a focus on the tactile feedback. However, it was acknowledged that this condition was the most difficult to play melodies with, and that it would take a longer time than was available in the study to take full advantage of it. This was confirmed by the results of the individual condition questionnaires where 'beat frequency' was consistently rated as the most mentally demanding.

> *It was best for fine tuning when I knew the approximate location of the note. It allowed me to shift completely into tune in these cases. The problem was with the melodic passages where the feedback is a bit too much to take in alongside performing. Works great for fine tuning of single notes but I wanted to turn it off otherwise.* P10
>
> *Very high level of precision, compared to all the other conditions* P8
>
> *Doing a glissando into a note helped a lot. Other ways of transitioning seem harder.* P7
>
> *Found it hard to distinguish between the 'resting' vibration of a note compared to the vibration when I was out of tune, especially as it varied between high and low notes according to their fundamental frequency.* P4

> *Similar to before, I had more to think about and think it might have had an adverse effect on other mental functions.* P1
>
> *I'd love to spend more time with the instrument to learn how to play it.* P5
>
> *It's very hard to fix quickly on the desired note.* P6
>
> *It felt the most organic of all the conditions, but this also made it very hard to perform the tasks at hand. Over a long period of time of playing the instrument, this one would probably be the most helpful as it provides a very high level of precision for tuning - but the learning curve is the hardest to begin with.* P9

## 4.6 DISCUSSION

The study examined how dynamic tactile feedback altered the navigation of pitch space, focusing on two main questions: first, how does dynamic tactile feedback impact pitch selection accuracy when compared to an audio-only condition; and second, what impact does dynamic tactile feedback have on the performer's actions and experience? Considering the first question, I found that the 'vibration when in tune' and 'beat frequency' conditions both provide improvements to tuning accuracy although with different temporal costs, as discussed in the following sections.

### 4.6.1 *Intonation*

The 'vibrations when out of tune' condition, although structurally similar to 'vibrations when in tune', was not observed to improve tuning performance across participants, suggesting that removing tactile stimulation for task confirmation rather than adding could impact feedback effectiveness. Although this condition did not improve tuning accuracy in comparison to the 'audio only' condition, participant preference for a particular *polarity* of tactile feedback highlights an important consideration when designing interfaces with tactile feedback: should the user be informed of successful or unsuccessful task completion? From this small sample we cannot conclude which is generally preferable, but it seems that such a guideline would need to be informed by both subjective user preference and objective measurement of performance.

On reflection it was noted that the experimental design of this study limited the statistical power of the findings. In order to get clearer comparisons of performer accuracy under the different conditions it may have been more appropriate to chose a smaller 'active' region of the touch sensor than the $\pm 8$ cents, which may have been too wide an error margin. A more fine-grained difference between conditions may have been observed in this case.

### 4.6.2 *The demands of real-time performance*

The improved accuracy in combination with the negative timing impact and performer survey results for the 'beat frequency' condition suggest that this tactile signal is either too nuanced or unfolds too quickly to be useful in a timing critical situation like the performance of a musical instrument, at least with the limited learning times that performers had in this study. Nevertheless, the fact that participants found the condition engaging as well as generally comprehensible means that this technique could possibly be fruitfully employed in a tactile interface in which fine accuracy is required but there are not the same time-pressures as a musical performance.

The reported reduced ability to maintain tempo when using vibrotactile feedback could be explained by the sensorimotor latency required to process haptic feedback and then act on it. It could be posited that expert musicians act in a *feed-forward* mode when performing, where planning and execution of musical passages happens at too high a rate to process the note-by-note responses from their instruments. Novice musicians on the other hand could be described as taking a *feed-back* approach where the sound and feel of each note are attended to. The position of these two processes in sensorimotor skill acquisition was discussed in Section 2.1.4.2 and Section 2.1.4.3 respectively. In the case of this study there was perhaps simply not enough time for the participants to advance to the feed-forward mode, to a point where the vibrotactile feedback was supporting their expectations rather than determining their note-by-note performance on the instrument.

### 4.6.3 *Gestural behaviour*

For the second question, I saw several emergent performance techniques that were directly influenced by the tactile feedback. The technique of lightly searching for a pitch before committing with greater pressure suggests that tactile feedback could be engaged immediately on contact with the instrument, with the sound only starting once the performer applies further pressure. This would allow a performer to confidently find the right pitch before producing a sound.

A series of emergent gestures that aided the performers in note-searching were also observed. Each of the feedback conditions influenced the performers' movements in different ways, encouraging varied temporal and spatial patterns of behaviour. This will be discussed in more detail in Section 4.6.5.

### 4.6.4 *Preference*

The high variability in results amongst performers precludes a statistically definitive answer to which condition is preferable; however, the alignment of accuracy and survey data yields support for the simpler cases, at least for situations where the user only has a short time to work with an interface. When choosing their overall preferred condition the vibrotactile feedback conditions were chosen over the audio only condition for all participants suggesting that they enjoyed having the additional signals present in the instrument.

Accounting for differences in preference and sensitivity to vibrotactile feedback is like accounting for variation in taste for music: there are great differences between participants and different interpretations. However, as mentioned in Section 2.3.1 in relation to Fontana et al. [91], I hope I have been able to raise important issues regarding haptic experience of dynamic vibrotactile feedback that hold beyond this small sample size.

### 4.6.5 *The projection*

I shall now discuss the observations from this study in relation to the projection model as presented in Section 3.6. This model considers the lens-like effect of an instrument-body combination which in this case is mostly defined by the capacitive touch sensor on the instrument's top surface: the pitch of the instrument's oscillator is controlled with the location of a fingertip on the sensor and velocity is controlled via touch area, approximately equivalent to touch pressure.

MUSIC THEORETICAL STRUCTURES Let us first question the role of the vibrotactile feedback in this instrument. At the very beginning of the experiment, when the instrument was played without any vibrotactile feedback, there was no implication that a diatonic scale was the desired mode to play in; the 2-octave pitch range of the instrument could be divided and traversed at will by the performer. As soon as the vibrotactile feedback was introduced it became clear that the continuous scale was demarcated into districts the size of either a tone or semitone. Each of the feedback conditions made a music theoretical structure present on the instrument [214]: the structure of a diatonic scale with fixed tuning was implanted upon the continuous 2-octave range of the instrument.

As the study progressed and the musicians realised that the diatonic scale was a reoccurring structure in the instrument and a requirement of the musical tasks they were taking part in, they interpreted the behaviour of the feedback in relation to this structure. In the case of 'vibrations when in tune', the vibrations were treated as a signal of confirmation, giving the performer affirmative feedback that they had reached an in-tune pitch. The vibrations of the instrument perform a type of quantisation of the pitch continuum. This feedback simplifies intonation on the instrument in a manner similar to the introduction of a keyboard to a stringed instrument (such as the nickelharpa) or to the difference between a fretless or fretted bass. In this case, however, the feedback does not restrict or force the performer towards an even-tempered scale, it rather provides a channel of feedback that a performer can use if they choose to do so, and if they have the mental bandwidth to do so. The audible pitch continuum of the instrument is not quantised itself, rather the feedback tries to guide the performer to play in accordance with the diatonic scale.

Let us consider the differences in the projections of the instrument with and without feedback. In both cases the actions that have meaningful effect on musical output are the same: the aperture of the instrument does not change between settings in terms of how action translates to sound. The difference lies in the way that the actions of the performer are structured by the characteristics of the feedback that they receive. We can imagine this as a musical theoretical structure being brought into focus by the projection between performer and instrument, influencing the emergent choreographies.

EMERGENT PLAYING TECHNIQUES In Section 4.5.4.2 the diversity of playing techniques that the performers employed in relation to different feedback conditions on the instrument was presented. With each feedback condition performers used subtly different actions related to the temporal unfolding of the vibrotactile feedback. The effects of each feedback condition on the performers' interpretation of the instrument can be seen in their resultant movements and the characteristics of their note-searching gestures. With the simpler feedback conditions there was an increase in the speed of performance in comparison to the condition with no vibrotactile feedback, whereas the more complex of the settings (beating) caused the performers to decrease the speed of their performance. Different means of arriving at a note were observed and this could be compared to the different styles of glissandi used when shifting between notes in string instrument playing. In this case however, instead of being used with a purely musical purpose in mind, it is rather a means of perceiving the vibrotactile feedback more accurately and then responding to it.

In Section 2.1.4.2 I discussed the *feedback* and *feedforward* modes of interaction that musicians employ when controlling an instrument in realtime. In the case of the above study we could posit that all the musicians remained too unfamiliar with the instrument to switch into a mode of performance that relies on prediction. Rather they were heavily dependent on the feedback the instrument produces. There were also individual cases of the feedback being detrimental to a performer's intonation, as they relied completely on the vibrotactile feedback which overrode auditory feedback.

The projection of action to sound remains ostensibly the same between settings on this instrument. What changes is the implied music theoretical structure of this projection. This is a characteristic of the lens, which becomes manifest through the vibrations of the instrument. Instead of forcing action to have a certain meaning this feedback could be understood as providing anchor points, landmarks, nodes and paths in the terrain of the instrument in a way parallel to the sonic output. These help structure a performer's mental representation of how action translates to sound and to how the instrument is laid out, aspects that are fundamental to learning a musical instrument (see Section 2.1.4).

## 4.7 CHAPTER CONCLUSIONS

This chapter has presented a study examining the impact of dynamic tactile feedback on the navigation of the pitch space on a self-contained digital musical instrument. The results suggest an improvement in pitch selection accuracy with certain types of vibrotactile feedback for the participants: in general accuracy was improved when participants performed with the 'vibrations when in tune' condition in comparison to the 'audio only' condition. An improvement in accuracy was also suggested in the 'beat frequency' condition however with a negative impact on timing and user response.

A set of emergent gestures linked to the type of feedback were also observed, suggesting that haptic information has a strong influence on how a performer conceptualises a new instrument. The statistical power of these results and their generalisability are necessarily limited to the particular instrument, study design and participants, but they point towards agreement with findings from previous studies, especially from Berdahl, Niemeyer, and Smith [26]. Notably, where previous studies used force feedback to push the performer to the right pitch, this instrument requires an active correction by the performer. Further investigations into the generalisability of these findings may have significant implications for the understanding of multimodal and cross-modal interaction design.

# 5

# AUDIO-HAPTIC ASYNCHRONY

*This chapter incorporates significant material from 'Action-sound Latency and the Perceived Quality of Digital Musical Instruments: Comparing Professional Percussionists and Amateur Musicians' by Jack, Mehrabi, Stockman, and McPherson, originally published in Music Perception 2018 [158] and 'Effect of latency on performer interaction and subjective quality assessment of a digital musical instrument' by Jack, Stockman, and McPherson, originally published in the proceedings of Audio Mostly 2016 [154]. The statistical analysis presented in Section 5.5 was conducted in collaboration with Adib Mehrabi.*

In this chapter I present an experiment that investigates the effects of small amounts of action-sound latency and jitter (10ms, 10±3ms, 20ms) on the interaction of musicians with a digital percussion instrument. I assess both the musicians' judgements of instrument quality and their timing accuracy under different latency conditions which are compared with a 0ms reference condition. Two groups of participants took part in this study, non-percussionist amateur musicians and professional percussionists. My aim with this research is to assess the impact of relatively small amounts of delay on the fluency and quality of the interaction, even when the auditory feedback is not perceived as detached from the action that produced it (the commonly accepted threshold for perceived audiotactile simultaneity can vary between 20ms and 70ms [275]). I also examine whether extensive rhythmical training and the demands of the musical task affect the influence of these delays on performance.

Whereas Chapter 4 focused on an instrument's dynamic feedback in response to a performer's actions, the work presented in this chapter takes a magnifying glass to the very instant of striking an instrument. The study presented in this chapter looks in detail at the temporal make-up of audio-haptic feedback while playing an instrument, and on how the alignment of stimuli between the two modalities can have subtle effects on how an instrument is experienced.

In Section 2.3 we discussed multisensory simultaneity perception and the time window within which multiple stimuli can be understood as correlated to the same action. The characteristics of the stimuli, their levels and perceptual relevance all adjust the size of this temporal window of integration. In the study that follows I have used a percussive DMI as a test bed for investigating the temporal relationship of haptic and auditory perception during interaction with an instrument and how this affects performers.

## 5.1 RELATED RESEARCH

Playing a musical instrument represents a highly-developed sensorimotor skill, where years of training and theoretical knowledge are brought together into the nuanced and expressive control required for musical performance. Delayed feedback (be it auditory, visual or tactile) can cause disruption to this sensorimotor control. Previous studies have shown that delayed auditory feedback (DAF) while performing with an instrument can disrupt musical production, primarily by increasing the variability of timing [290, 293, 386]. This disruption varies with the length of the delay; similar effects have been shown for DAF in speech [147].

In the field of HCI delayed feedback has mostly been studied as system latency (the asynchrony between a control gesture and a system's corresponding response) and jitter (the variability of this asynchrony). Latency is a fundamental and unavoidable issue affecting interactive digital systems and their linkages between virtual and physical worlds, and has long been recognised as potentially harmful to a user's experience of control [210, 243]: even if accuracy of temporal performance is not impacted, the *qualitative* experience of the user may be negatively impacted [171]. The way latency and jitter affect a user have been shown to vary greatly depending on the specific demands of the task and situation [7], for example in the case of direct or indirect touch, or of tapping or swiping on a touchscreen [169, 261].

In the present study, a group of highly trained professional percussionists and a group of non-percussionists (with varying levels of musical experience) evaluated the effects of variable levels of DAF on a novel digital percussion instrument. My aim is to investigate the differences in the effects of these delays on timing accuracy and perceived instrument quality between the two groups, and to understand the influence, if any, that specialised training in rhythm-based musical practices has on these measures.

### 5.1.1 *Feedback delays in HCI*

The *effects* of asynchronies within multisensory feedback have been extensively studied in the field of HCI. Levels of acceptable latency and jitter have evolved with the capabilities of the technologies underpinning digital systems, and the more accurate measurements they allow. In the late 1960s a measure of 100ms between touch input (via a stylus) and visual response was considered adequate for a system to be described as responsive [249]. Recent research has shown this amount to be much too large, redefining the meaning of responsiveness in HCI, and has demonstrated that the limits of acceptable latency and its effects on user experience vary widely depending on the specifics of the task and the combination of modalities involved [172].

Due to the current proliferation of touchscreen technologies much effort has focused on measuring acceptable levels of latency in such devices[1]: when examining multisensory latency in touchscreen buttons Kaaresoja, Brewster, and Lantz [173] suggest that latency should be lowest for the tactile channel (5-50ms), followed by audio (20-70ms) and finally the visual (30-85ms). For direct-touch systems (where feedback is presented beneath the fingertips) the threshold of noticeable visual-tactile latency (with visual delayed) has been shown to be as low as 69ms for tapping and 11ms for dragging [69]. These findings point to the task dependency of latency perception. Annett et al. [7] have conducted studies with stylus interaction which revealed that the ability to perceive latency is worse whilst being involved in a focused task like drawing or writing (approx. 50 ms), in comparison to perception whilst performing arbitrary tasks (scribbling, dragging) reported previously as (approx. 2-7ms) [262]. Even if no delay is perceived latency can still impact the quality of interaction: Ng et al. [261] have shown that in the case of delayed visual feedback in a direct-touch system participants greatly prefer lower latencies, and even those well below 10ms when no delay is perceived.

### 5.1.2 *Delayed auditory feedback in acoustic musical instruments*

Auditory delays within an acoustic musical context are multifaceted and commonplace. Lago and Kon [190] point out the variability of the effects of such delays and their dependence on instrument type, style of music and spatial positioning: in ensemble playing, for example, delays ranging from 10ms to 40ms can often be present due to the distance between players and the speed of sound [56]. Perceptible amounts of latency between action and sound do not normally affect acoustic instruments, which produce sound in reaction to action instantaneously, as the sound producing mechanism and control interface are one and the same: the only factors introducing latency here are the action of the instrument and time that it takes for the sound to travel from the point of excitation to the performer's ear. There are however some exceptions where latency is built into the mechanism of an instrument – in the case of a piano, the delay between a key reaching the key bottom and the hammer striking the string can be around 35ms for *pp* notes and -5ms for *ff* notes. These figures do not include the key travel time (the time elapsed between initial touch and the key reaching the key bottom) which for pressed

1 Current touch systems typically take 50-200ms to update the display in response to a physical touch action [261]

touch can be greater than 100ms for *pp* notes and 25ms for *ff* notes [11].

Lester and Boley [200] provide a comprehensive overview of the effects of latency during a direct live sound monitoring situation in a recording studio and in which they found that sensitivity to latency was highly dependent on instrument type and monitoring style (in-ear versus wedge monitors). As latency increased it became less of a spectrum-altering phenomenon (comb filtering due to the combination of acoustic and the slightly delayed monitoring feed) and more of a temporal perception issue (above 6.5ms caused temporal smearing with certain instruments for in-ear monitoring). The low thresholds found in their paper are in part due to the specifics of the live monitoring situation where acoustic and delayed sound are combined.

### 5.1.3 *Delayed auditory feedback in DMIs*

The close coupling of action to sound via the virtual mechanism of the computer is of prime importance to building compelling DMIs. There are many parts of a DMI that introduce latency and jitter between action and sound: buffering in hardware and software, latency in the audio code itself (from frequency domain processing for example), transmission delay between sensors and audio engine due to USB connection, latency induced by smoothing or signal conditioning of the sensor input [239, 382]. These factors can combine to create a significant delay between performer action and resultant sound and impede what Wessel and Wright describe as the development of *control intimacy* between performer and instrument [374] as we discussed at length in Section 3.3.

Wessel and Wright [374] suggest that DMIs should aim for a latency of less than 10ms with less than 1ms jitter. In an evaluation of common techniques used to create DMIs conducted by McPherson, Jack, and Moro [239], it was demonstrated that Wessel and Wright's guideline is still not met by many commonly employed toolkits. Although this guideline is widely respected in the field it is also based on intuition and experience rather than on empirical studies. This is a point that the study presented in this chapter hopes to clarify, bringing together empirical studies conducted in music psychology, HCI and psychophysics to focus on the specific case of the DMI.

## 5.2 RESEARCH QUESTIONS AND SCOPE

This study addresses Research Question 2:

*How important is audio-haptic simultaneity, that is the delay between action and sound, to the perceived quality of a DMI, and how does this vary* *with musical expertise?*

I address this question by conducting a comparative experiment with a percussive DMI with variable levels of action-sound latency. This experiment was conducted with amateur musicians and professional percussionists to assess the impact of expertise on latency perception. In particular I address the following sub-questions of RQ2:

(a) At what point is action-sound latency perceptible in a DMI?

(b) What influence does action-sound latency have on the perceived quality of a DMI, in particular in relation to tangible qualities?

(c) What influence does action-sound latency have on rhythmic performance on a DMI?

(d) How do two groups of performers with different levels of expertise compare in terms of the above questions?

These questions are answered by assessing the sychronisation error of each performer from each expertise group under each one of four latency conditions. Each of the latency conditions are directly compared to a zero latency case and rated for a series of instrument quality measures. This is combined with a thematic analysis of structured interviews conducted with each of the performers and an analysis of gestures employed.

## 5.3 instrument design

### 5.3.1 *Measuring multimodal delays*

The challenge of measuring multimodal delays in interactive systems has been explored in many studies, mostly oriented towards touch-screen interactions [172, 325, 384]. In the present study, in order to counter the common problems of latency in a DMI and to have sufficient control over the exact amount of latency and jitter in the system I have again used the Bela platform[2] (see Section 1.5) as the basis of the instrument. Bela provides sub-millisecond latency and almost jitter-free synchronisation (within 25$\mu$s) of audio and sensor data [239, 240] making it a suitable platform for controlling the exact amount of latency present in a DMI.

### 5.3.2 *Physical design*

The technology probe (see Section 3.5.4) I created for this experiment was a self-contained percussive digital musical instrument, the playing surface of which consists of eight ceramic tiles of varying sizes

2 http://bela.io/

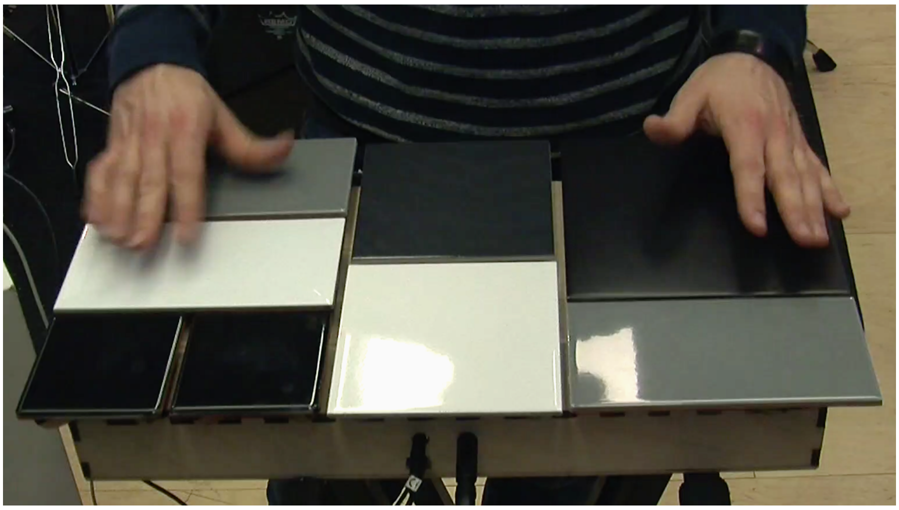


Figure 5.1: The instrument built from eight ceramic tiles with piezo discs attached to them to sense vibrations.

(see Figure 5.1). The instrument represents a 'simplest case' digital percussion instrument with one dimension of control: discrete velocity triggering of samples. Each of the ceramic tiles has a piezo disc vibration sensor attached to the back with pliable scotch mounting tape. Mounts for the tiles were created from laser cut plywood and felt, which suspend the tiles by their antinodes allowing them to vibrate freely when struck thus ensuring a strong signal to the vibration sensor [331]. A layer of 3 millimetre rubber foam was glued to the back of each tile to further condition the signal while attenuating the acoustic resonance of the tile.

The piezo sensors are connected via a voltage biasing circuit to the Bela board. A striking action on the tiles induces vibration in the tile which is passed through signal conditioning routines and a peak detection algorithm, detailed in the next section, before being used to trigger samples of Gamelan percussion instruments. The intensity of the strike is measured when a peak is detected and mapped to the amplitude of the sample playback.

#### 5.3.2.1 *Filter group delay and peak detection*

The peak detection routine includes a DC blocking filter, full-wave rectification and a moving average filter. Strikes were detected by looking for peaks in the sensor readings using an algorithm that looks for an increase in the reading followed by a downward trend once the reading is above a minimum threshold. When a peak is detected the amplitude of the strike is measured and then assigned to the sample appropriate to the tile. The synthesis engine had enough computational power to play 40 simultaneous samples, and I used an oldestout voice stealing algorithm if all voices became allocated, to allow for fast repeated strikes.

The audio output on Bela used a sample rate of 44.1kHz and a buffer size of 8 samples. The analog inputs used for the piezo discs were sampled at 22.05kHz, synchronously with the audio clock. The total action-sound latency consists of the duration of the two buffers (360$\mu$s) plus the conversion latency of the sigma-delta audio codec (430$\mu$s). In addition to this there is the group filter delay of the FIR filter (moving average) that was used to smooth the piezo signal over 20 samples before the peak detection of 1 sample, resulting in 250$\mu$s delay. As the analog inputs and audio outputs are synchronised on an individual sample level, jitter between them is no more than 25$\mu$s [239]. In the present study the sound of the instrument was monitored directly through noise-cancelling headphones. I conducted a test of the headphones to ensure that the noise-cancelling function was not introducing additional latency and found that when the noise-cancelling was turned on there was an additional 100$\mu$s latency in comparison to the analog signal path. I take this total of 1.2ms latency as Condition A, which I call the "zero latency" condition because the distance between the tiles and the ears would normally contribute around 2ms of acoustic latency due to the speed of sound.

### 5.3.3 *Sound mapping*

The use of a piezo vibration sensor naturally gives us an ergotic link [207], as discussed in Section 3.3, between the force of a strike and the amplitude of the sound output. The response curve of this sensor is linear, unlike other commonly used sensors in electronic percussion instrument like FSRs. For a full review of sensors commonly used in percussion instruments see the work of Medeiros and Wanderley [242] and Tindale et al. [352]. By using this sensor I was able to naturally preserve the relationship between physical energy at the input and perceived physical energy at the output by producing a linear software relationship between input level and output level.

In the present study four sample sets are used, each consisting of eight individual samples that were assigned to each of the eight ceramic tiles. Gamelan samples were chosen due to the variety of different striking types in this instrument family, and to further reinforce the ergotic link between percussive strike and percussive sound behaviour. All samples have equal duration, equal pitch variation and perceptual attack time. The four sample sets can be further divided into two groups characterised by the perceptual acoustic features of their attack transients. The difference between these two groups is in the spectral centroid during the initial strike: they can be broadly described as brilliant sounding or dull sounding (striking a metallic bar with a metallic beater versus striking a metallic bar with a padded beater). Pitch height was preserved on the instrument for each sample set moving from left to right, with the lowest pitched note mapped to the largest tile on the left-hand side, and the highest pitched note mapped to the smallest tile of the right hand side, increasing vertically for each third of the instrument (see Figure 5.1).

The peak detection and triggering routine remained constant throughout the experiment while the latency condition and sample set changed. Throughout the experiment the raw signals from the instrument were recorded onto an SD card by Bela for later analysis. This included the signal from each of the eight piezo disks attached to the tiles, the audio output, and the metronome or backing track that the participants were monitoring through headphones in the second and third tasks.

## 5.4 STUDY DESIGN

### 5.4.1 *Method*

This study was conducted in a sound-isolated studio. The instrument was mounted on a keyboard stand whose height the participants could adjust for comfort. On a podium next to the instrument there was a laptop where the participants input their responses and changed the settings of the instrument. Participants monitored the instrument directly through noise cancelling headphones (Bose QC-25). White noise was played in the room through a PA system at (50dB) at a level where all acoustic sound from the instrument was inaudible when the participant was performing. This was to avoid participants hearing any excess sound coming through air conduction from their contact with the instrument, focusing their attention on the sound that was presented through the headphones and their haptic experience of the strike.

### 5.4.2 *Participants*

Two groups of participants took part in the study I am presenting. The first group (referred to as "non-percussionists" or "NP" from now on) consisted of 11 participants (8 male / 3 female) whose age was between 26 and 45 years and who were recruited from my university department. All members of this group had musical experience but none were professional. 8 of the 11 participants classified themselves as instrumentalists and the other 3 as electronic musicians. None of this group had received training in percussion. These participants had varying degrees of musical training (0-15 years; M = 9.2 , SD

= 4.5 ). All but 2 of the participants had used a computer to make music, with 6 of the participants regularly using the combination of a hardware controller and software instrument to compose and/or perform music.

The second group (referred to as "professional percussionists" or "PP" from now on) consisted of 10 participants (9 male / 1 female) whose age was between 26 and 35 years. They had completed at least a Bachelors degree in performance specialising in percussion and were working professionally, either as performers in orchestras, as session musicians or in education. This group had between 10 and 20 years of formal percussion training (M = 13.8, SD = 2.5). All participants in this group had received training on a second instrument (2-15 years; M = 11, SD = 3.5) most commonly piano in the case of 6 participants. Both groups reported normal hearing and normal or corrected-to-normal vision. This experiment met ethics standards according to my university's ethics board. Professional percussionists were paid £30 and amateur musicians £10 for taking part in this experiment.

### 5.4.3 *Experimental set-up*

Four variable latency and jitter conditions were tested:

- Condition A: 'zero' latency
- Condition B: 10ms latency
- Condition C: 20ms latency
- Condition D: 10ms latency $\pm$ 3ms latency (simulated jitter)

These conditions were created by delaying the sound triggered by a detected strike by a set number of samples and were verified on an oscilloscope. In the jitter condition each strike was assigned a random latency between 7ms and 13ms. I chose these three specific latency conditions based on a recent series of measurements conducted by McPherson, Jack, and Moro [239] of common techniques used to create digital musical instruments. This study found that amounts of latency above 10ms where common place in such tool chains, with jitter often above 3ms. I also deliberately chose the maximum latency condition (20ms) to be within the thresholds of simultaneity perception for audio-haptic stimuli as found by Adelstein et al. [2] of around 24ms. This was to focus my findings on the effects of latency and jitter when a delay is not necessarily perceived between action and sound. These four conditions were chosen to give us a large enough sample of comparisons between each condition.

### 5.4.4 *Procedure*

The study lasted for approximately one hour and 15 minutes, and consisted of two sections followed by a structured interview. Participants were video and audio recorded throughout the experiment.

#### 5.4.4.1 *Part 1: Quality assessment*

In order to evaluate the participant's subjective impression of quality of the different latency conditions on the instrument, I decided to use a method that involved participants rating the conditions in comparison to one another for a series of quality attributes. In this part of the study latency conditions B, C and D were always compared to condition A. This part of the experiment was inspired by Fontana et al.'s study on the subjective evaluation of vibrotactile cues on a keyboard [91]. In their study the impact of different vibrotactile feedback routines on the perceived quality of a digital piano is assessed, and my methodology and analysis in Part 1 takes a similar route. In this first section participants were presented with the instrument and advised to freely improvise while switching between two settings, $\alpha$ and $\beta$. Their task, for each pair of $\alpha$ and $\beta$, was to comparatively rate the two settings according to four quality metrics (Responsiveness, Naturalness, Temporal Control, General Preference) drawn from studies on subjective quality assessments of acoustic instruments [317] as discussed in Section 3.5.3 and based on the qualities I hypothesised would be most relevant to the changing latency conditions. Once they had decided on the comparative ratings of the two settings they then moved onto the next pair.

STIMULI AND CONDITIONS Between $\alpha$ and $\beta$ I changed both the latency condition and sample set. I deliberately wanted to mask the changing latency conditions to evaluate whether the latency conditions were perceivable by the participants when they were not instructed to focus on the amount of latency present. When starting the study, participants were instructed simply to compare the different settings on the instrument according to the attributes and to try and not base their ratings on preference for a sample set alone: the fact that latency would be present and changing was not mentioned.

EXPERIMENT PROCEDURE The instrument was self-contained, dealing with all the sensor and audio processing via Bela allowing participants to monitor the instrument directly via noise cancelling headphones. To switch between $\alpha$ and $\beta$ a separate laptop was used, which hosted a GUI built in PD[3] which communicated with the Bela board

3 https://puredata.info/

Figure 5.2: Experimental set-up with instrument and accompanying laptop for changing settings.

via User Datagram Protocol (UDP), allowing participants to switch between settings at will (see Figure 5.2). For each pair the zero latency condition (A) was assigned to either $\alpha$ or $\beta$ in a weighted random order, while the other setting in the pair would always contain a latency condition (B, C or D). Two different sample sets were also selected in a weighted random order for $\alpha$ and $\beta$. There were 12 such pairs, again presented in a weighted random order, for each participant, ensuring that each sample set was in the zero latency position 3 times per participant. This meant that each participant rated each pair of latency conditions 4 times, each time with a different sample set assigned to each of the conditions. Participants were advised to take around 35 minutes to complete the evaluation of the 12 pairs.

Participants also input their ratings on the accompanying laptop via a GUI that consisted of slider inputs for each attribute using a Continuous Category Rating (CCR) scale, a rating widely used in subjective quality assessments of interfaces. While rating the settings, participants were instructed to improvise freely with no restrictions on their chosen style. Participants moved the slider on the continuous scale to rate the relative merits of the two settings (see Figure 5.3). The scale had the following titles along its scale:

- $\alpha$ is much better than $\beta$
- Both $\alpha$ and $\beta$ are equal
- $\beta$ is much better than $\alpha$

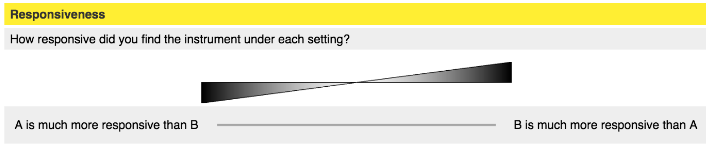



Figure 5.3: Continuous input slider for rating the settings in comparison to one another.

#### 5.4.4.2 *Part 2: timing accuracy*

In order to evaluate the impact of the latency conditions on the temporal performance of the participants I used a synchronisation task in which they were instructed to play along with a metronome under each latency condition. A metronome at 120 bpm was played through the headphones. The participant was then instructed to tap along with the beat using a single tile, dividing the metronome beat into progressively smaller chunks: every crotchet (quarter note) which is equivalent to the 120 bpm of the metronome, then every quaver (eighth note), then every semiquaver (sixteenth note). They performed each of these tapping exercises for at least four bars, paused and then moved onto the next. They repeated the whole task three times for each latency condition and then moved onto the next condition. Each of the four latency conditions were presented in a weighted random order and the sample set remained the same across participants. My methodology in this part of the study was derived from Fujii et al.'s study on synchronisation of drum kit playing [94].

#### 5.4.4.3 *Part 3: structured interview*

To conclude the experiment a structured interview lasting between 10 and 20 minutes was conducted. The interview was conducted in front of the instrument and demonstrations were encouraged from the participants. The following themes were discussed in each case: general impression of the instrument, including the styles of playing that worked well or did not; techniques used to distinguish between $\alpha$ and $\beta$ in Part 1, the free improvisation; whether they noticed what was changing between settings, besides sample set; their experience of latency as an issue in musical performance.

### 5.4.5 *Data collection*

Alongside the survey results I also collected performance data from the instrument: the audio output and vibration signals from each tile. The audio streams were sampled at 44.1kHz while the vibration signals were sampled at a rate of 22.05kHz. Audio and video were recorded throughout.

## 5.5 FINDINGS

### 5.5.1 *Quality judgements*

In this section the analysis of perceived quality focuses only on task 1, the comparison of settings.

#### 5.5.1.1 *Statistics*

In the analysis and in Figure 5.4 condition A (zero latency) is always $\alpha$ (zero on the y axis) for legibility, although in the study it was randomly assigned to either $\alpha$ or $\beta$. For each group (professional percussionist, non-percussionist) we fitted separate Linear Mixed Effect Regression (LMER) models with fixed effects of quality (responsiveness, naturalness, temporal control, general preference) and condition (10ms, 20ms, 10ms±3ms), and random intercepts for each participant. The models were fitted using the `lme4` [17] package for R [301]. We conducted a full factorial Type III ANOVA on each LMER model, with Satterthwaites's degrees of freedom approximation from the `lmerTest` package [187].

#### 5.5.1.2 *Group 1: non-percussionists*

Figure 5.4 (a) shows the median and IQR for all participants in this group, (c) the mean and standard error. On average condition A (the zero latency condition) was rated more positively for all qualities than condition C and D, the 20ms and 10ms±3ms conditions. We found a significant effect of condition ($F(2, 517) = 7.4, p < 0.001$). A Tukey post-hoc analysis on each factor shows that the effect of condition is driven by a significant difference between 10ms±3ms and 10ms ($Z = 3.4, p_{adj} < 0.01$) and 20ms and 10ms ($Z = 3.2, p_{adj} < 0.01$) (all p-values were adjusted using the Benjamini and Hochberg false discovery rate correction (FDR=5%)).

#### 5.5.1.3 *Group 2: professional percussionists*

Figure 5.4 (b) shows the median and IQR for all participants in this group, (d) the mean and standard error. For the professionals, we found significant effects of condition ($F(2, 470) = 4.9, p < 0.01$) and quality ($F(3, 470) = 5.0, p < 0.01$). A Tukey post-hoc analysis on each factor shows that the effect of condition is driven by a significant difference between 10ms±3ms and 10ms ($Z = 3.1, p_{adj} < 0.01$) and a significant difference between 20ms and 10ms ($Z = 2.5, p_{adj} < 0.05$) (all p-values were adjusted using the Benjamini and Hochberg false discovery rate correction (FDR=5%)).

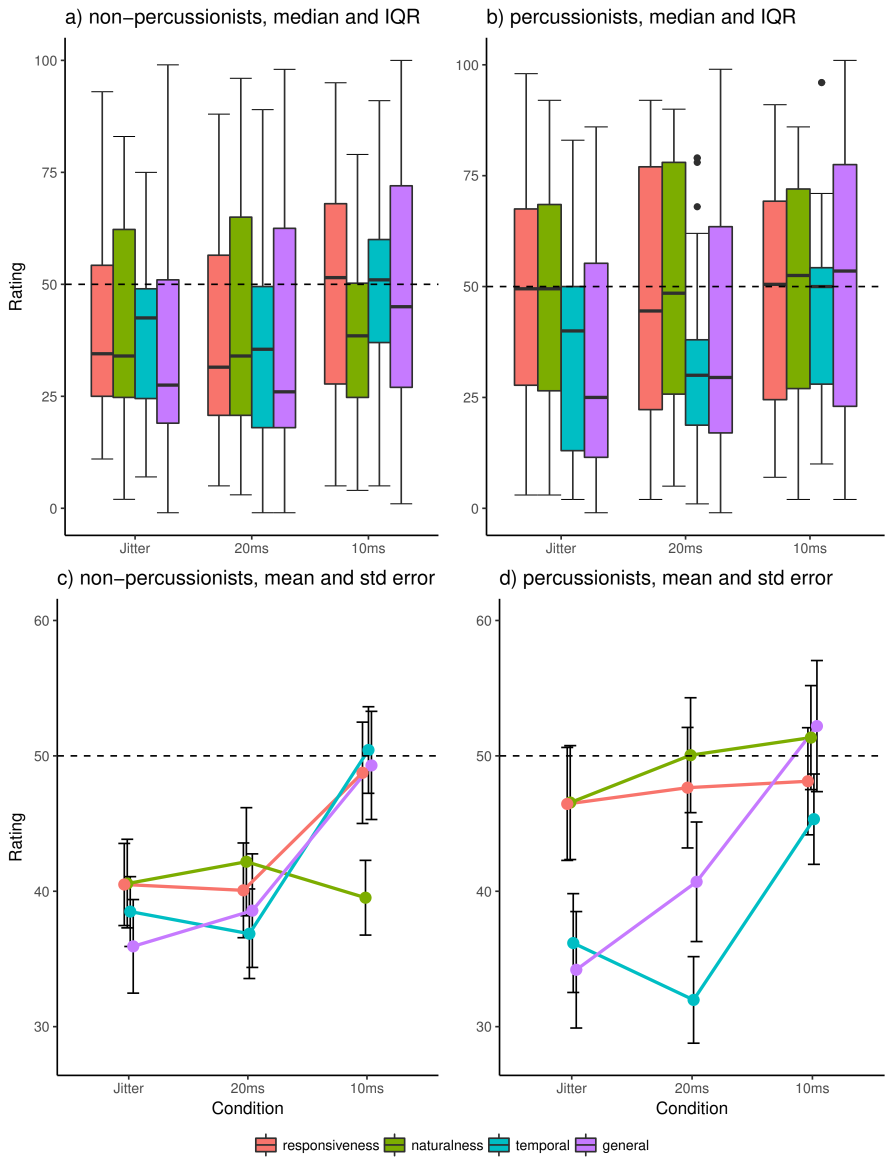



Figure 5.4: (a) and (b) show the median and IQR of all responses from both participant groups. (c) and (d) show the mean and standard error of all responses from both participant groups. 0 on the y axis corresponds to '$\alpha$ is much better than $\beta$', 100 with '$\beta$ is much better than $\alpha$', and 50 with 'both $\alpha$ and $\beta$ are equal'. Note that in this representation 0 always means that condition A (zero latency) is preferred to the other latency condition it is being compared to.

INFLUENCE OF SAMPLE SET For both groups we tested to ensure that sample set was not having an overriding effect on quality ratings (i.e. that participants were basing their ratings on sample set alone). When fitting the LMER models we also included sample set as a fixed effect and found no significant effect, so were able to discount this as a factor.

### 5.5.2 *Temporal performance*

In this section the analysis of timing performance focuses on task 2, playing with a metronome. For this analysis we compared the onset of the strike against the onset of the metronome tone, looking for the difference between the timing of the strike on the tile and the metronome tone, rather than the audio output of the instrument, which had added latency under three conditions. The onset of each strike relative to that of the metronome was defined as the synchronisation error (SE). The value was negative when the onset of the strike preceded that of the metronome and positive when the strike onset lagged behind the metronome.

#### 5.5.2.1 *Statistics*

For the modelling we fitted an LMER model with fixed effects of group (non-percussionists, professional percussionists), temporal division (crotchet, quaver, semiquaver) and condition (10ms, 20ms, 10ms±3ms), and random intercepts for each participant. As with the quality judgement analysis, the significance of each fixed effect was tested using a full factorial Type III ANOVA on the LMER model, with Satterthwaites's degrees of freedom approximation.

TYPICAL DISTRIBUTION Figure 5.5 shows the typical distribution of strikes of both groups for each tempo measure and each latency condition. Figure 5.6 presents the median and interquartile range (IQR) for all latency conditions for both groups. For the NP group we excluded one participant from the analysis due to them having a Mean Synchronisation Error (MSE) of 30% greater than the group MSE giving 10 participants in this group. All 10 participants in the PP group had a MSE within this threshold.

#### 5.5.2.2 *Synchronisation error*

Figure 5.6 shows the median and IQR for all participants under each latency condition and division. We found a significant effect of condition ($F(3, 2289) = 5.9, p < 0.001$) and division ($F(2, 2289) = 3.7, p < 0.05$). A Tukey post-hoc analysis showed that the effect of condi-

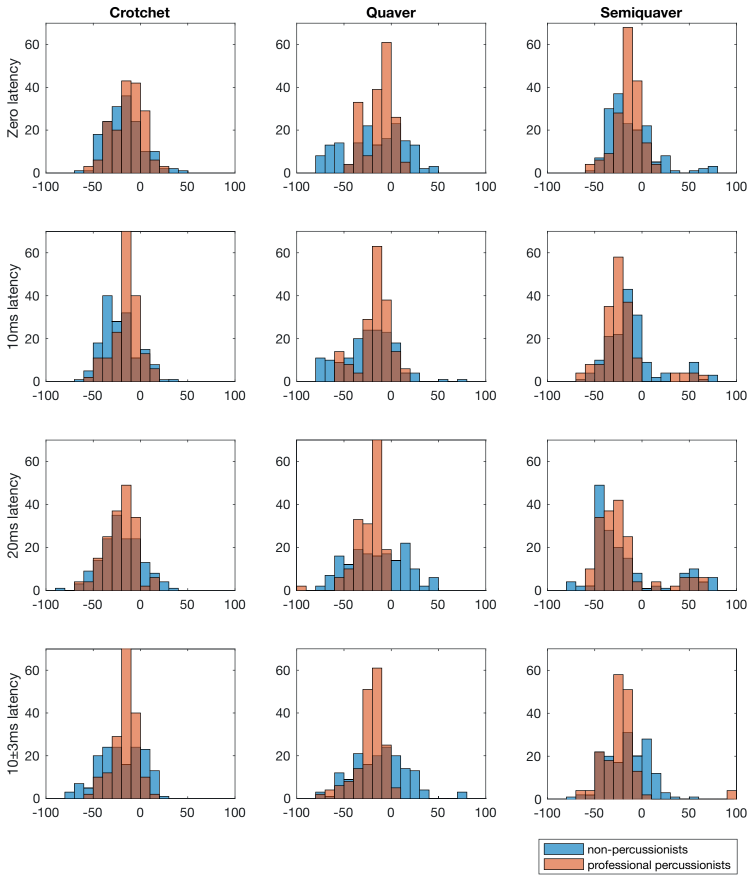



Figure 5.5: Distribution of strikes for both groups showing the spread of the timing of their strikes during the synchronisation task.

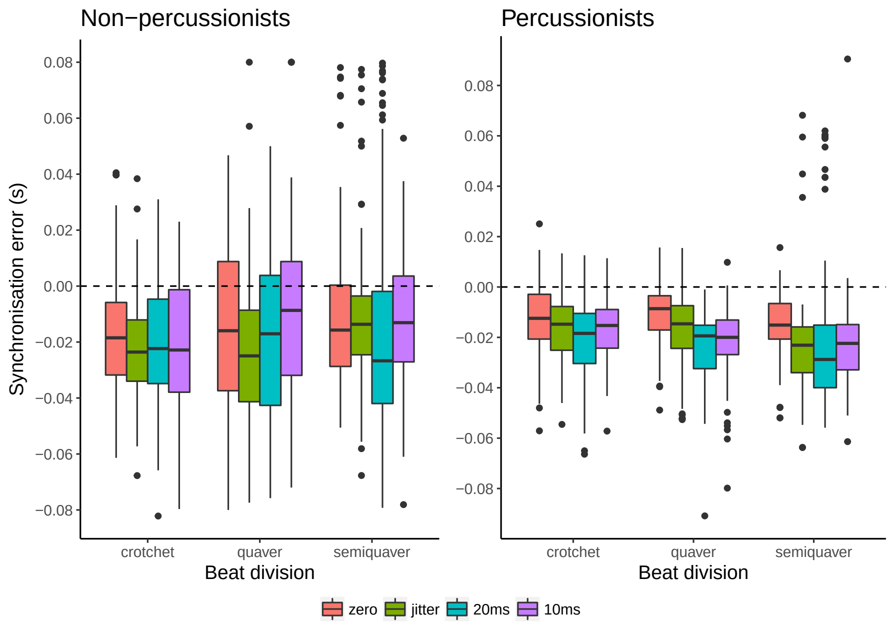



Figure 5.6: Median and IQRs of synchronisation error for the first rhythmic task for all latency conditions for both groups.

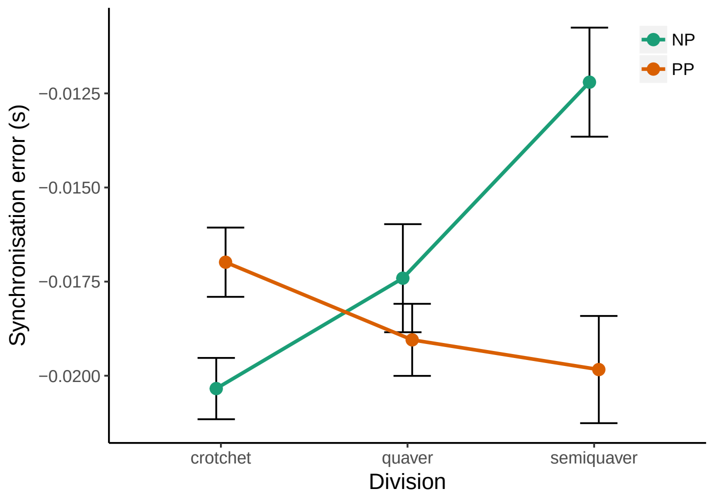



Figure 5.7: Interaction contrasts between division and group for synchronisation error.

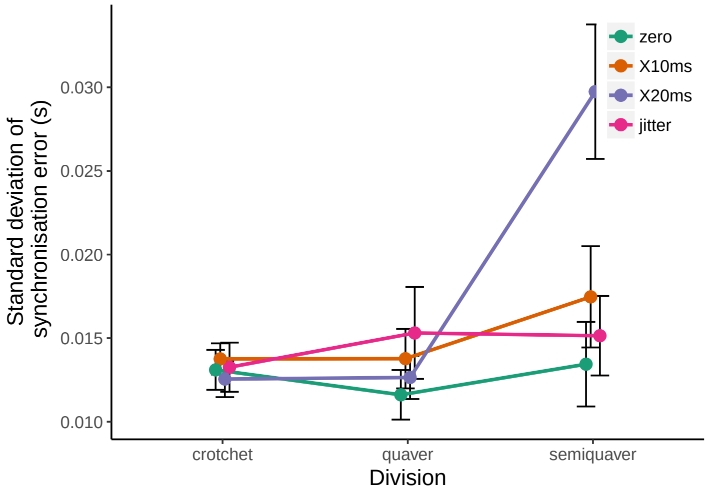



Figure 5.8: Interaction contrasts between condition and division for the standard deviation of synchronisation error.

tion is driven by a significant difference between 10ms and zero latency ($Z = 3.6, p_{adj} < 0.001$) and between 20ms and zero latency ($Z = 3.7, p_{adj} < 0.001$). A smaller and marginally significant difference was seen between 10ms±3ms and zero latency ($p_{adj} = 0.08$). We also tested for interactions between each fixed effect (by fitting new models with interaction terms), and found a significant interaction between group and division ($F(2, 2288) = 12.5, p < 0.001$). This effect is shown in Figure 5.7, where synchronisation error is negatively correlated with IOI for the NP group, and the opposite effect is observed for PP. A post-hoc analysis of the interaction contrasts between all factors of group and condition was conducted using the `phia` package [67] for R. This showed a significant interaction between group and all 3 division factors: *crotchet-quaver* ($\chi^2(1) = 4.4$, $p < 0.05$), *crotchet-semiquaver* ($\chi^2(1) = 25$, $p < 0.001$), and *quaver-semiquaver* ($\chi^2(1) = 7.8$, $p < 0.01$).

To assess the effect of division for each group, we refitted separate models for each group, with fixed effects of condition and division, and random intercepts for each participant. In the case of the non-percussionists this showed a significant effect of division ($F(2, 1483) = 17.3, p < 0.001$) and no significant effect of condition. A Tukey post-hoc analysis showed that the effect of division is driven by a significant difference between *crotchet-semiquaver* ($Z = 5.9, p_{adj} < 0.001$) and between *quaver-semiquaver* ($Z = 3.7, p_{adj} < 0.001$). In the case of the professional percussionists this showed a significant effect of condition ($F(3, 806) = 8.8, p < 0.001$) and no significant effect of division. A Tukey post-hoc analysis showed that the effect of division is, as expected from the interactions above, driven by a significant difference between 10ms and zero latency ($Z = 3.0, p_{adj} < 0.01$), 20ms and zero latency ($Z = 4.6, p_{adj} < 0.001$) and 10ms±3ms and zero latency ($Z = 4.3, p_{adj} < 0.001$). This suggests that the effects of condition on timing accuracy for all participants found in the initial tests were driven by the results from the PP group.

#### 5.5.2.3 *Variability of synchronisation error*

To evaluate the variability of timing accuracy we refitted the above mentioned model but with standard deviation of the synchronisation error as the dependent variable. We observed heteroskedasticity in the residuals of the fitted model, which we rectified using a log transform of the dependent variable (sderror). We found a significant effect of group ($F(1, 20) = 16.6, p < 0.001$), division ($F(2, 220) = 5.5, p < 0.01$), condition ($F(3, 220) = 4.2, p < 0.01$) and an interaction between condition and division ($F(6, 220) = 3.5, p < 0.01$). The mean standard deviation between groups (across all conditions and divisions) is 0.018 for non-percussionists and 0.013 for percussionists: this is a difference of over almost 50%.

Upon testing for interaction contrasts between condition and division we found the significant interactions are between 20ms and each of the crotchet-semiquaver ($\chi^2(1) = 13.2$, $p < 0.01$), and *quaver-semiquaver* ($\chi^2(1) = 8.4$, $p < 0.05$) conditions. This can be seen in Figure 5.8. We noted a medium but non-significant positive correlation between error and standard deviation of error (i.e. as error decreases, so does its variation).

### 5.5.3 *Interviews*

The structured interviews conducted at the end of the study were annotated and then coded using a thematic analysis framework [40]. The coding strategy aimed to identify the major themes that related to latency perception and judgements of instrument quality. Other themes that came from these interviews related to style, the constraints of the instrument and the evolution of gesture over the duration of the study shall be discussed at length in Chapter 7.

#### 5.5.3.1 *Non-percussionists*

AWARENESS OF LATENCY Latency perception was the first theme I investigated: whether the settings with latency were perceived as having a delay or not. Only 3 out of the 11 participants stated that there was latency or delay changing between the settings. This suggests that either the amounts of latency were small enough to not be perceived as a delay for 8 of the 11 participants or that the changing sample sets masked the changing latency conditions. When asked what was changing between settings aside from the sample set participants generally reported a changing responsiveness and level of dynamic control: they described shifting triggering thresholds at times, that the instrument was catching less of their strikes under certain settings, or that the dynamic range of the instrument was expanding and contracting, factors that were not in fact changing. In the quality ratings from Part 1 of the study I saw the zero latency condition receiving more positive ratings than the 10ms$\pm$3ms and 20ms latency condition for the attribute Temporal Control. This suggests that a disruption to the temporal behaviour of the instrument was identified even if its cause was not established as delayed auditory feedback. Some participants also acknowledged that under certain conditions they were struggling to maintain timing although, again, they did not specifically identify that a delay or latency was the cause:

> *...one was very difficult to keep some sort of stable timing on, while the other one just clicked for some reason and made a lot more sense.* P4
>
> *On the second one* (condition A) *I didn't have to put much thought into it or didn't have to tap myself in or anything. It was just there under my finger tips.* P10
>
> *...I was playing very fast passages and seeing if it captures all the notes. In some of the settings it wasn't tracking well but in others it was.* P8

These quotes point towards the complexity of the 'response' of the instrument: this term does not seem to have been reduced to how fast the instrument responded but rather is about how much of the participant's playing was translated into sound by the instrument: judgements seem to be based on how participants felt the instrument was reflecting the energy they put in. In addition to the above reports, there were also additional multimodal effects of the latency conditions reported, where the perceived effort required to play a note increased with latency.

**REPORTED EFFECTS OF LATENCY DURING THE STUDY** 4 of the 11 participants reported that under certain latency conditions they felt they needed to strike the instrument with more force to get the instrument to respond in the way they wanted.

> *I also noticed that I had to put more energy into one or other of the pairs to get a sound from the instrument.* P9

For these four participants I analysed the variation in striking velocity across latency conditions to test if their reports of increased force of strike was influenced by latency condition. I found that for these participants there was indeed an increased mean velocity of strike for 20ms and 10$\pm$3ms latency in comparison to the zero latency condition [154].

#### 5.5.3.2 *Professional percussionists*

AWARENESS OF LATENCY In general the professional percussionists were more aware of the latency conditions than the non-percussionists, with 9 out of 10 mentioning it as the changing factor between settings.

> *I felt like some of them were a bit more 'on top'* [...] *sometimes you felt like it wasn't instantaneous and you're not connected to it.* P4
>
> *The latency also changed as well and they* [the sample set] *weren't necessarily related.* P1
>
> *...sometimes there was a bit of a delay, sometimes the note was behind the strike.* P3

Professional percussionists were also more conscious of latency as an issue that faces digital musical instruments. In some cases this came from their experience of using digital samplers in live performance or from experiences of home recording with a backing track.

> *I have a set of TD Roland drums, and they have latency, it's very slight but I definitely notice it, it's more than it would be from an acoustic kit definitely.* P5

Many of the participants also explicitly mentioned latency as a negative factor in an instrument's design that impedes their performance.

> *...with percussionists, we're so used to, you hit it and, bang, it's there. So any kind of delay is a bit disconcerting.* P7
>
> *When it sits on top it's a lot more enjoyable to play. I know when that happens you tend to forget that you're playing something, and you tend to explore, you make music then, rather than trying to work out the instrument.* P4

ABILITY TO ADJUST TO ACTION-SOUND LATENCY Participants also spoke of their ability to adjust to the changing latency conditions naturally and without too much active thought when they were freely improvising without a metronome.

> *Because I've got experience of adjusting I was able to adjust to what I was hearing. I do that naturally. When playing acoustic instruments you listen to what's coming out and you adjust to it. It's always a tiny little difference, you can adjust naturally. You deal with it. You can get by.* P3
>
> *We do have experience with working with delay and trying to think about that. You need to compensate so that you don't sound late. You don't really think about it much normally, it's too much if you think about it, has to be by feel.* P7
>
> *I was adjusting very quickly. If I was doing them all with a click track I think the response I gave would be different as I would*

> *actually feel myself trying to adjust to where the beat was when there was latency, like this I just did it without thinking.* P5

experience with acoustic instruments From the PP group there were many comparisons made between latency in a digital instrument and the timing adjustments that orchestral percussionists have to do as they switch their position in the orchestra or switch the acoustic instrument they're playing. Talk of 'sitting behind the beat', 'sitting in front of the beat' and 'sitting right on the beat' described how the percussionists conceptualise the micro-adjustments they make to their timing in order to ensure that the conductor (and audience) hears them as in time with the rest of the ensemble. When asked how they manage to adjust their playing like this, most stated that they had no idea how they actually did it: it was something that they had learned from being told by a conductor or other performers that they were coming in early or late, and at this point in their careers it was just a necessary part of their role that they were able to do without thinking. They mentioned that the dress rehearsal before a concert was the most important in terms of making this adjustment, as their timing needs to be tuned to their position in the ensemble and the acoustics of the room.

> *If you're sitting at the back of the orchestra the physical sound getting to the front takes longer as you're so far back. And for certain instruments this can take even longer. A lot of the time you have to play a little bit ahead or behind the beat to make sure it fits with everything else.* P3

Percussionists must also adjust to the mechanical action of the instrument they are playing. Professional percussionists are multi-instrumentalists that are expected to master and be able to switch between many different instruments in a matter of seconds. This brings with it the ability to switch playing techniques quickly and to adjust playing style to the specific action of an instrument – what the percussionists referred to as an instrument sounding *early* or *late*. Tambourine was given as an example of an instrument that sounds late, as was tubular bells and timpani. Examples of instruments that speak early were given as triangle and other metallic instruments played with hard beaters. The notion of how an instrument speaks seems to be related to frequency range of the instrument but also to surface hardness, the action of the instrument (triangle versus church bell for example), striking type (hard versus soft beaters, played with the hands or not), although conflicting examples were given by different percussionists.

> *In the case of this instrument* [the instrument used in this study] *it's to do with the fact that it's hard. I know that hard surfaces sound immediately, whereas floppy surfaces sound later,*

> *like timpani. I guess that's just sort of Pavlovian – it's hard, it's going to sound quickly.* P7

## 5.6 DISCUSSION

### 5.6.1 *Quality judgements*

In terms of the comparison of latency conditions and the quality judgements participants gave via the survey, there were trends of agreement between both groups. The results from Part 1 suggest that latency of 20ms and 10ms±3ms can degrade the perceived quality of an instrument in terms of temporal control and general preference, even when the amount of latency is too small to be perceived as a delay by the performer. This is in agreement with findings from Kaaresoja, Anttila, and Hoggan [171] when evaluating the impact of audio-tactile latency on user interaction with touch-screens. The fact that condition D (10ms±3ms latency) was rated in a similarly negative manner as condition C (20ms latency) in relation to the zero latency condition, but that condition B (10ms latency) did not receive similar negative ratings, highlights the importance of stable as well as low latency. This points to an agreement with Wessel and Wright's recommendations [374] of 10ms latency with 1ms of jitter as a goal for digital musical instruments.

None of the participants in this experiment performed with a mean degree of accuracy in Part 2 that was better than the jitter amount (±3ms), yet this condition was still rated negatively. This suggests that subtle variation in the stability of the temporal response of an instrument can be detected by performers even if they cannot perform with a degree of accuracy that is less than the jitter amount. These findings, alongside previous work [308] suggest that the amount of acceptable latency and jitter does not correspond directly to the limits of sensorimotor accuracy possible by the player.

### 5.6.2 *Latency perception*

In the non-percussionist group 3 of the 11 participants identified latency, or delay, as the changing factor between settings. For the other 8 participants the difference between settings was reported as a changing triggering threshold or dynamic range, both of which remained identical throughout the study. There were also reports of the feel of the instrument changing and the effort taken to play the instrument increasing with conditions where there was latency.

In the professional percussionist, group 9 of the 10 participants reported latency or delay as the changing factor between settings. It seems that this group was much more aware of latency and its causes from their experience as orchestral players, and were generally better at discussing it, as can be seen in the quotes presented from the structured interview. Making micro-adjustments to the timing of their performance in relation to an ensemble or to their instrument is a common part of professional percussionist's role as a musician. This may explain the difference between the groups: professional percussionists display superior synchronisation ability [51] and timing acuity [77], whether from their extensive training on a rhythm-based instrument or from natural ability.

### 5.6.3 *Timing accuracy*

#### 5.6.3.1 *Effect of latency and beat division on mean synchronisation error*

The findings from this study suggest that both groups timing accuracy was impacted by latency condition and division in different ways. For the NP group it was found that increasing division of the beat was affecting the accuracy of their playing, whereas latency condition showed no significant effect: it was observed that the MSE and variation of MSE increased as the beat division increased. This suggests that the error in their temporal performance increased as they were required to strike faster. No significant effect of latency condition on timing accuracy was found for this group.

The opposite appears to be the case for the PP group: timing accuracy was significantly affected by latency condition and not by beat division in all but the *semiquaver* case. For this group the zero latency condition had a significantly lower MSE in comparison to the other three latency conditions across beat divisions. The standard deviation of MSE under each latency condition did not differ significantly for each beat division aside from for 20ms latency at the smallest beat division (semiquaver) as can be seen in Figure 5.8. This suggests that for the larger divisions this group did not find latency disruptive to timing accuracy: when they were required to play at a speed above a certain threshold (IOI of 125ms) the latency became detrimental to their performance. Our findings suggest this was mostly the case with the 20ms latency condition, which would equate to 16% of the IOI of 125ms when playing semiquavers at 120bpm, well above the variation in timing accuracy from professional percussionists that has been previously reported [66]. It is worth noting that with the PP group the 10ms condition also appears to have measurably degraded their timing accuracy, particularly when performing at higher speeds. This is one of the few noticeable differences between 0ms and 10ms, suggesting that with the higher rhythmic acuity of the percussionists even 10ms can have a negative impact on their performance.

#### 5.6.3.2 *Mean synchronisation error*

A generally higher degree of variation in the MSE of the NP group was observed in comparison to the PP group, as can be seen in Figure 5.6. This agrees with the findings of Manning and Schutz [222] that participants with high levels of rhythm-based training (particularly percussionists) show superior timing abilities (MSE and variability of MSE) and temporal acuity in comparison to other musicians and non-musicians.

The group means of the MSE for the zero latency condition and across all metronome divisions ranged from -17 to -7ms for NP and from -15 to -12ms for PP. The mean standard deviation ranged from 20 to 33ms for NP and 8 to 12ms for PP. The MSE and SD for the NP group were larger than that found by Fujii et al. in their study with highly trained percussionists [94] where a mean synchronisation error of -13 to 10ms was achieved for a metronome with standard deviations of 10 to 16ms, whereas the MSE and SD of MSE for the PP are roughly aligned with these findings. The MSEs of both groups were smaller than those reported in previous tapping studies with non-musicians in which MSE was usually around -20 to -80ms, while for the NP group they were roughly equivalent to the performance of amateur musicians -10 to 30ms [9, 307]. Several of the participants in the NP group did have a degree of musical training, just not specifically percussion training, which could explain this finding. The values of MSE for the PP group in this study were smaller when compared with the finger tapping study of Gerard and Rosenfeld [108] who found an MSE of -25ms in professional percussionists.

A further analysis step that falls beyond the scope of this thesis would be to investigate systematic synchronisation errors in the performances of each of the groups. In this respect part of the synchronisation error that was observed could be attributed to systematic and reoccurring time deviances [137].

#### 5.6.3.3 *Adaptation and negative mean asynchrony*

When tapping along to a metronome participants commonly exhibit a Negative Mean Asynchrony (NMA): they anticipate the beat and strike early by between 30 and 10ms (see Aschersleben [9], Repp [306], and Repp and Doggett [307]). A measure of NMA and the variability of this asynchrony are a common way of assessing the temporal accuracy of a performer [308]. The bi-directional influence of auditory and movement information is evident in many simple tapping studies where auditory information guides motor timing: audiotactile stimuli with delays to auditory feedback cause anticipations to increase with the amount of delay [10] when delay is gradually introduced up to 70ms in a tapping test, whereas NMA reduces with deafferented participants with only auditory and visual feedback [343].

This measure can be significantly affected by adaptation to asynchrony (see Vroomen and Keetels [366] for a review). The adaptation process is typically evaluated by measuring participants' perceptions of crossmodal simultaneity both before and after an exposure period, where there is a constant feedback delay between the stimuli presented in the two modalities. Vroomen and Keetels [366] describe this as a widening of the temporal window for multisensory integration. The temporal window for audiotactile integration has been shown to widen in response to a relatively short exposure to asynchronously presented tactile and auditory stimuli in the case of passive tactile perception [260].

In the PP group an increase in negative mean asynchrony during the crotchet and quaver beat divisions that partially reflected the amount of latency being added to the instrument was also observed. This was an increase in MSE of approximately 5ms and 10ms for the 10ms and 20ms latency condition respectively. This can be seen quite clearly in Figure 5.6. It seems that there was a degree of compensation in relation to the latency condition but it was not an anticipation of the full latency amount (i.e. moving a strike 20ms earlier when 20ms of latency was present to bring the auditory feedback in line with the metronome). Anticipation of strike to match sound has been observed by others when introducing larger amounts of delay to auditory feedback [10, 66, 343]. These anticipation effects were not observed with the NP group for any latency condition. It could also be hypothesised that as a result of the experimental method, where the amount of latency was changed regularly between conditions, the adaptation as reported in other studies did not have enough time to occur [366].

### 5.6.4 *Rhythmic training and latency perception*

Regardless of training, participants generally agreed on judgements of perceived instrument quality. In the case of the non-percussionists even if the latency was not perceived as a delay, its effect on the fluency of interaction with the instrument appears to be recognised by the participants. Timing accuracy in the non-percussionist group was not significantly affected by the latency condition, yet this group rated 20ms and 10ms$\pm$3ms latency negatively in comparison to the zero latency condition. From the structured interviews they were reports that certain conditions felt 'under the fingers', whereas with others the connection between action and sound was not as clear. This highlights the subtlety of the effects of latency and the specific demands of percussion instruments where sound is a result of direct unmediated interaction.

In general the PP group were much more aware of latency and able to identify it as the changing factor between settings and talk explicitly about adjusting for latency. This is perhaps due to the extensive rhythmical training they have undertaken and their expertise in switching between instruments with different actions. Some of the percussionists spoke about the changing latency conditions as the changing action of the instrument: whether the instrument would sound 'late' or their playing would be 'right on top' of the beat allowing them to forget the instrument and concentrate on making music. This connects with ideas of instrument *transparency* as discussed in Section 3.3.3: Nijs, Lesaffre, and Leman [264] propose musical instruments as mediators between gesture and sound output. Transparency in this mediation is the point where the performer doesn't need to focus attention on the individual operations of manipulating an instrument, instead focusing on higher-level musical intentions. Latency in this interaction might be understood as a barrier to transparency.

Latency perception and the effects of latency vary widely dependent on the nature of the musical task, style of playing, instrument and individual experience of the performer. From this study it is not possible to conclude what the acceptable amount of latency that DMIs should aim for in general is; also, as the sample size is relatively small, there needs to be a degree of caution in interpreting the results, as their statistical power is necessarily limited by the amount of participants in each group. My aim with this study is to highlight the effects of small amounts of latency on the perceived quality of an instrument, an effect that I propose as similar to the degradation of feelings of presence in VR situations: latency as "a cause for reduction of suspension of disbelief" [4]. In the case of DMIs the notion of 'presence' is perhaps best equated to the ergotic aspects of an instrument: how energy is maintained in the digital system in the translation of action to sound [207].

### 5.6.5 *The projection*

In terms of the projection model we can imagine latency as effecting the transparency of the lens reponsible for the projection, that is the degree to which action passes to sound undistorted. Latency may prevent certain choreographies developing on an instrument, for example those that require a high degree of temporal control or simply high tempo performance. A deeper exploration of the instrument used in this study as it encounters the professional percussionists is presented in Chapter 7.

## 5.7 CHAPTER CONCLUSIONS

This chapter has presented a study investigating the impact of latency and jitter on the temporal accuracy of performance and judgements of instrument quality for two groups of participants; professional percussionists and non-percussionists (with varying amounts of musical experience). The experiment involved quality assessments of a novel percussive instrument with variable latency (zero, 10ms, 10ms±3ms, 20ms), temporal accuracy tests and structured interviews.

My findings relating to judgements of instrument quality reveal that both groups expressed a preference for zero latency in comparison to 10ms±3ms and 20ms latency. Importantly, the zero latency and 10ms latency conditions show no significant difference in rating for either group. This suggests that a stable latency of 10ms is acceptable to performers of a DMI, whereas 20ms is not. The 10ms±3ms latency condition was rated in a similarly negative manner to 20ms when compared to the zero latency condition, suggesting that the addition of a random jitter of ±3ms is enough to negatively affect the perceived quality of an instrument. The results support the recommendation put forward by Wessel and Wright [374] that DMI designers should aim for a latency of 10ms or below with a jitter of 1ms or less, however my findings cannot tell us exactly what the minimum threshold of acceptable latency is, except that it must be somewhere between 10ms and 20ms.

The ability to perceive latency varied between groups, as did the impact on temporal performance. Generally, professional percussionists were more aware of the latency conditions and better able to adjust for them in their playing, although this ability decreased as the temporal demands of the task increased. For the PP group the 10ms latency condition measurably degraded their timing accuracy when performing at high speed. This is one of the few noticeable differences between 0ms and 10ms, without which it would be tempting to conclude that there is no reason to aim for latency under 10ms. This study shows that even though 0ms and 10ms latency might be rated qualitatively similarly, participants with a high degree of rhythmic training still perform less accurately with 10ms latency in comparison to 0ms, suggesting that this level of latency does in fact matter.

Latency can degrade the illusion of action translating to sound, a factor that is central to expressive and skilled control of DMIs. In this study I have demonstrated the effects of latency on two different groups of musicians and found marked differences between each group in terms of disruption to timing accuracy and in terms of their ability to identify latency. Both groups were in agreement as to the impact of latency on the quality of the instrument in question. Findings from this study suggests that the influence of latency on the perceived quality of a digital system does not hinge on the temporal acuity of the user, rather it is something that can degrade the fluency of the interaction regardless of skill.

Importantly for this thesis there were a number of effects of action-sound latency that could be described as tangible in nature. The feel of the instrument changed with latency: the perceived effort required to produce a note, the perceived weight of a tile, whether the instrument 'sounded immediately' or 'sounded late', and other factors related to the perceived *action* of the instrument. These reports shall be discussed further in Section 8.2 and highlight how action-sound latency in a DMI can have tangible effects.

# 6 CONTROL INTIMACY, RICHNESS AND PHYSICAL SUPPORTS

*This chapter incorporates significant material from 'Democratising DMIs: the Relationship of Expertise and Control Intimacy' by Jack, Harrison, Morreale, and McPherson, originally published in the proceedings of the International Conference on New Interfaces for Musical Expression, NIME 2018 [159] and 'When is a Guitar not a Guitar? Cultural Form, Input Modality and Expertise' by Harrison, Jack, Morreale, and McPherson, also published in the proceedings of NIME 2018 [130]. The study presented in this chapter was conducted collaboratively with Jacob Harrison.*

In the previous two chapters we discussed active haptic feedback in an instrument and the temporal composition of audio-haptic feedback respectively. Each of the studies evaluated how these tangible design parameters affect the experience and performance of differently skilled groups of performers. This final practical study brings the focus to *control intimacy* and to the *richness* of an interaction. This can be broadly explained as the degree to which a performer's physical actions when playing an instrument are translated into the resultant sound. The experiment presented in this chapter uses four guitar-derivative DMIs to explore the influence of different levels of control intimacy, and of variations in tangible design elements, on the performance and experience of two differently skilled groups of performers.

## 6.1 RELATED RESEARCH

Chapter 3 discussed how interfaces often fail to take advantage of the sensorimotor capabilities of a performer: rich control gestures are often transformed into an approximate representation of that gesture via the sensing strategy and its mapping onto the control of a sound engine, as a trajectory or series of discrete events. This is not a problem in cases when instruments are designed to take agency over sound production and eschew sensorimotor control; however it does affect cases where a close coupling between action and sound is desired.

Section 3.3 discussed the various parameters that can affect the intimacy of control that a performer has over an instrument. One of these parameters was the temporal coupling of action and sound, explored in the study presented Chapter 5. Another key aspect of control intimacy is the resolution with which a gesture is captured by the computer, and how the chain of mediation within the instrument maintains the nuance and micro-details of a performer's movements. In the case of an instrument with decreased control intimacy (for example with high temporal and spatial quantisation and dynamic range compression) the musical results can remain highly compelling, however the finer detail and nuance of a performer's movements is removed from this interaction. In this case, interaction becomes primarily constrained by the technological capabilities of the instrument rather than the sensorimotor capabilities of the performer.

Within NIME and related research fields, much discussion has centred around richness of control, the level of detail of control that a performer has over an instrument. It has been proposed for many years that richer mappings between control input and sound output are better [374], and much design effort has gone into this idea. Yet in practice, the design of DMIs involves balancing two factors that can often seem at odds with one another: the steepness of the learning curve that a performer has to climb to make music with an instrument, and control intimacy, how the richness and nuance of a performer's movement translates into musical output. In the experiment presented in this chapter the aim is to interrogate the idea that a richer instrument is a better instrument.

As discussed in Section 2.4, embodied music cognition offers a theory of music-making and reception that is deeply rooted in body movement. From the embodied perspective music is based on action, and musical meaning is tied to movement characteristics and patterns: the presence of bodily motion, whether that body is real, apparent or absent in the music [289], is at the basis of music cognition. Considering this theoretical angle, this study sought to investigate the influence that varying levels of control intimacy, that is the degree to which a performer's movement is reflected in an instrument's sound, have on interaction and on judgements of instrument quality.

Sensorimotor 'familiarity' is another essential aspect of instrumental control. Learning a musical instrument is as much a haptic as it is an aural cultural practice, and learning how to move the body in certain defined ways is an integral part of musical training. As discussed in Chapter 2, the development of an internal model of how action translates to sound and of how bodily movements are used to control an instrument depends on the formation of sensorimotor pathways that couple the biomechanics of the body to the mechanics of the instrument. Tangible aspects of an instrument's design are of prime importance to the formation of this internal model: they shape the movement of a performer and act as landmarks, anchor points, and guides for trajectories. In their response they become emblematic of the instrument's character and behaviour, for example in the case of strings, keys or valves.

The experiment presented in this chapter investigates the influence of specific tangible design cues on a performer's experience of a DMI. The physical form of an instrument plays a role ergonomically, in the way that it forces the body to contort and conform to play it, but it also has great symbolic power: the form of an instrument carries a wealth of cultural associations and social rules beyond its ergonomic function. In this case we were interested in looking at design cues that function on two different levels in how they influence musical movement. The first relates to the *macro* sound-motion features [120] as discussed in Chapter 2, and was modulated in terms of the global form of the instrument. The second relates to the *meso* and *micro* sound-motion features and is explored through variable input modalities. Whereas the previous two studies presented in this thesis used a single physical instrument with variable settings that altered its behaviour, this study is based on four separate instruments that try to tease apart tangible design cues, allowing us to make comparisons between them. This study was conducted in collaboration with Jacob Harrison, whose research focuses on accessibility in musical instrument design. Although in conducting the study we were answering separate research questions, and have made separate contributions to the field, we designed the experiment and the instruments it utilises in collaboration, and thus much of the text that follows is presented in the first person plural.

## 6.2 RESEARCH QUESTIONS AND SCOPE

This study addresses Research Question 3:

*How does control intimacy (the degree to which a performer's actions are reflected in the behaviour of an instrument) affect the perceived quality of a DMI, and how does this vary with musical experience?*

I addressed this question by conducting a further comparative experiment with a series of guitar-derivative DMIs that varied in levels of control intimacy (audio-driven synthesis vs. sample triggering), as well as in global physical form (guitar shaped vs. tabletop) and input modality (strings vs. touch sensor). Two groups with varying musical expertise (non-musicians and guitarists) took part. In particular this study addresses the following sub-questions of RQ3:

(a) What influence does the level of control intimacy have on judgements of instrument quality and gestural behaviour?
(b) What influence do physical form and input modality have on performer experience?
(c) How important is the reinforcement of interaction metaphors through tangible guides in an instrument's design?

(d) How does expertise influence the above questions?

These questions are answered by comparing the responses of each group to a pair of guitar-derivative DMIs with variable control intimacy, with which a series of free improvisations and musical tasks were completed. I reinforce findings from the participant responses through gesture observation and through the thematic analysis of structured interviews.

## 6.3 INSTRUMENT DESIGN

In this study we attempted to create a scenario in which we were able to break down an instrument into a series of tangible design cues that could act in a modular fashion and be interchangeable between instruments. In order to do so, we took one base instrument design inspired by the guitar and utilised physical modelling to simulate the strings. What followed was an iterative design process during which, by creating various prototype instruments, we endeavoured to draw apart some of the design elements that are normally closely related in an instrument's design. The tangible design elements investigated in this study were not dynamic factors that changed in relation to performer action as in the previous two studies, but rather related to physical features of the instrument's design: its physical shape and support structures for the hands of the performer while playing.

The instruments were four guitar-derivative DMIs with varying combinations of overall form-factor (guitar-shaped vs. tabletop) and interaction modality (plucked strings vs. touch sensor) which can be seen in Figure 6.1 and are herein referred to as SG (Strings-Guitar), ST (Strings-Tabletop), TT (Touch-Tabletop) and TG (Touch-Guitar). The two stringed instruments also featured a switch that allowed the richness of the transfer of action to sound to be altered (audio-driven synthesis to sample-triggering). The instruments were created using the Bela platform [240] (see Section 1.5) which was used for the sensing and string modelling to create a high-performance, low latency embedded instrument in each case. Again each of the instruments was created as a technology probe [150], as discussed in Section 3.5.4.

### 6.3.1 *Physical construction*

This was by far the easiest factor to split apart, as it only required altering the global housing of the instrument mechanism without having great implications on the technology used inside. As the instrument's sound quality was not dependent on the resonance or on the material characteristics of the body we were able to freely variate the global form until we found two contrasting forms that satisfied our design purposes.


Figure 6.1: The four instruments designed for this study. Clockwise from top left: Strings-Guitar (SG), Strings-Tabletop (ST), Touch-Guitar (TG) and Touch-Tabletop (TT)

GUITAR FORM: Both instruments were designed to be played using similar techniques to a guitar or other strummed string instruments. A single enclosure for the 'Guitar' instrument was built by Ailish Underwood, a model maker from Bournemouth University of the Arts who collaborated on the design side of this project. The enclosure is constructed out of hardwood and has a sculpted neck with six push-buttons roughly at the position of the lower frets of a standard size guitar. The buttons are arranged in two columns of three, set to chords I, IV and V in the key of G on the top column, with their relative minors on the column below. The body contains a cavity, allowing the sensing method to be changed between the string module or touch sensor module.

TABLETOP FORM: Two similar enclosures for the tabletop instruments were designed, one of which can be seen in Figure 6.2, and were intended to reflect design cues from boutique music hardware, such as standalone synthesisers. The push-buttons were placed in the lower left-hand corner, with the strumming area placed at a 45 degree angle, which was found to be a comfortable method of strumming on a tabletop.

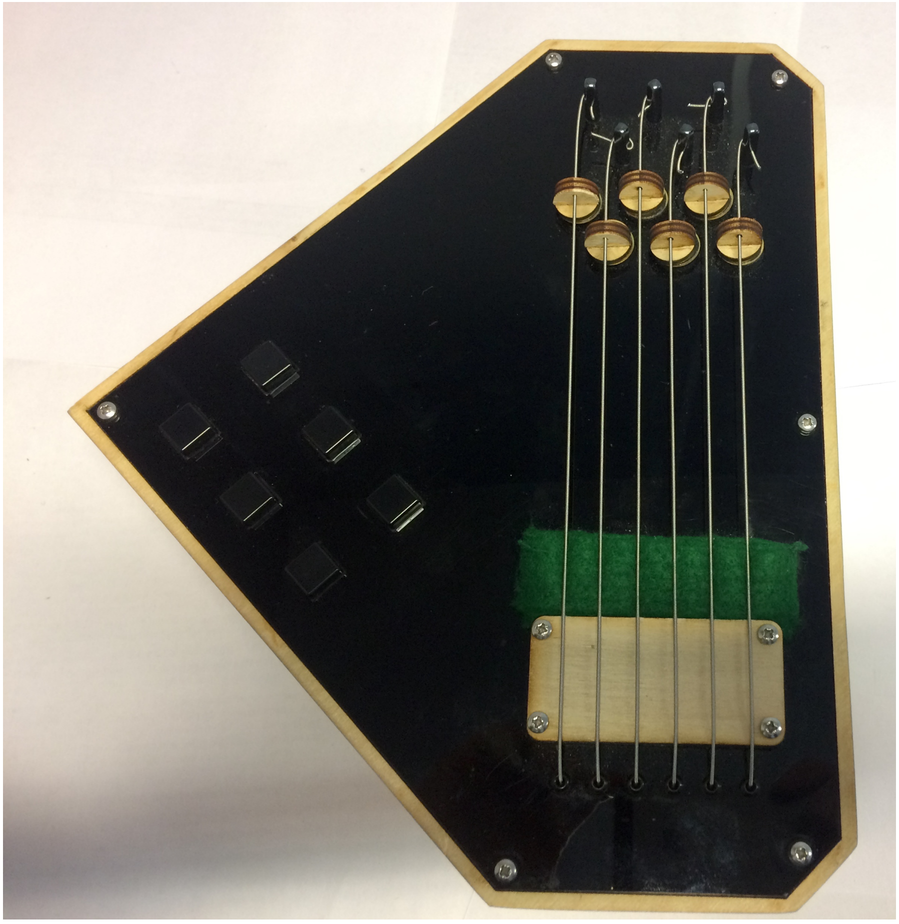


Figure 6.2: The strings tabletop instrument with piezo bridges under each string.

Figure 6.3: Close-up of the piezo bridge in the string versions of the instruments.

### 6.3.2 *Sensing Techniques*

In order to variate the input modality, we used two different sensor technologies, which would both allow for a similar type of control, namely strumming and finger-picking. We chose two sensing strategies that would represent familiar forms of interaction for the participants. The physical strings with vibration sensing would be immediately familiar for the guitarists, as the instrument shared many characteristics with string instruments. The capacitive touch sensor would be familiar to both groups from the kind of interaction that is common with touchscreens.

physical strings: The string instruments have six short lengths of .040 gauge bass guitar string held over a 'strummable area' of about 10cm. At one end, the strings are terminated over a block of felt-covered foam; at the other end they terminated with six individual bridge-pieces at the other with integrated piezo disc sensors, and held to a low tension using adjustable zither pins. This provides an acoustic impulse similar to the attack of a plucked string on a guitar. The thick strings and low tension produce a short decay and fewer resonant properties than a typical guitar string held to tension. The details of the piezo bridge set-up can be seen in Figure 6.3.

capacitive touch sensor: The touch instruments use a capacitive touch slider derived from the TouchKeys keyboard overlay design [238] to detect finger position, the same sensor used in the instrument presented in Chapter 4. Six 'string areas' are equally spaced along the sensor. Several layers of paint were applied in thin strips to the surface of the sensor, to provide tactile cues as to the location of each string area as can be seen in Figure 6.4. This type of sensor was chosen for its 'swiping' and 'tapping' affordances, gestures commonly associated with touchscreen interfaces but which have a direct analogy to strumming and finger-picking.

### 6.3.3 *Sensor mappings*

In order to have a variable level of control intimacy we had to focus on a sensing strategy that would allow us to have a rich form of interaction, similar to that of an acoustic instrument, that could then be simplified. Two varying levels of intimacy were implemented on the string input modality, as with this sensing technique we were able to alter the amount of richness and nuance of control gesture making it through to the sound output. This was achieved by comparing audio-rate sensing, where the raw signal from the strings was excit-

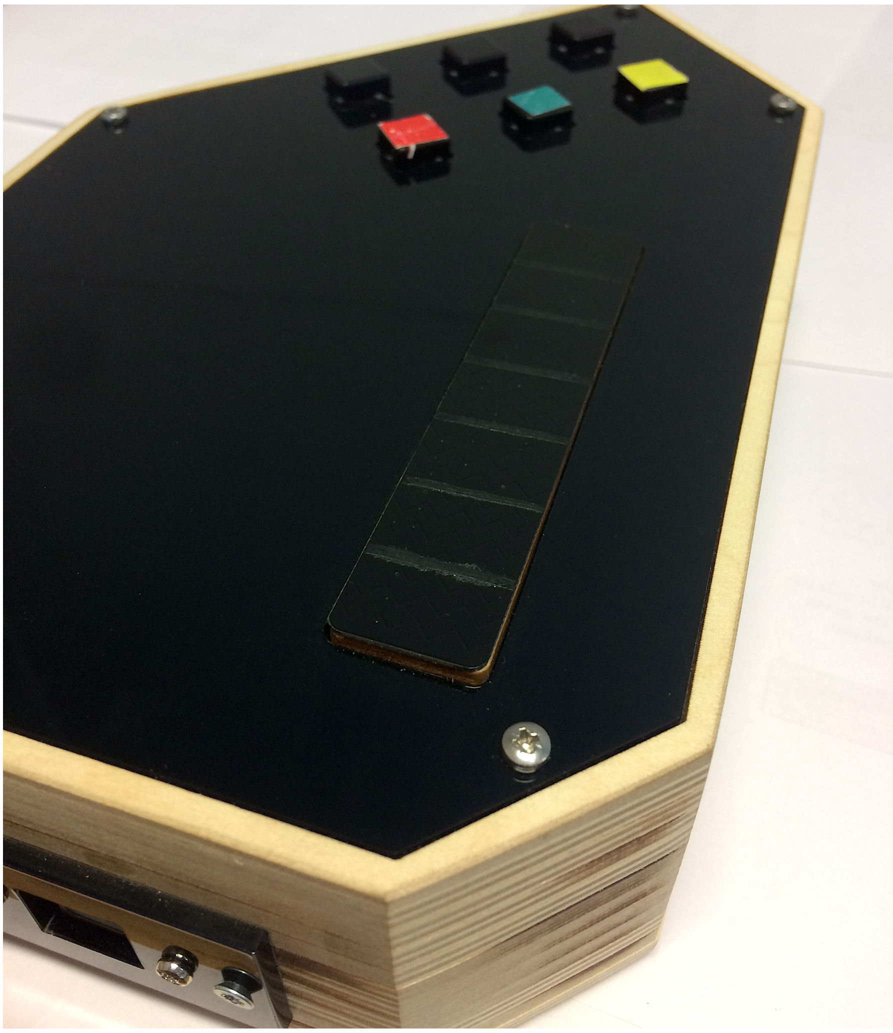


Figure 6.4: Close-up of the capacitive touch sensor version of the instrument with tactile ridges across the sensor to represent string placement.

ing a physical model, with sample-triggering, where playback of the recorded impulse excited a physical model.

At the basis of all four instruments was an implementation of the Karplus-Strong plucked string algorithm [160] which was used to simulate six virtual strings which were excited in real time using signals from the piezo and touch sensors. The physical model can be excited by an audio input.

STRINGS WITH AUDIO-RATE EXCITATION: Excitation of a virtual string model using a real-time audio signal has been implemented and documented in previous NIME research, including the Kalichord [323], BladeAxe and PlateAxe [248] and Caress instruments [251]. Such instruments allow intuitive control over the resulting sound by varying the way the virtual strings are excited (plucking hard or soft, or with different materials). The instruments in this experiment follow a similar principle, but feature dampened strings terminated over piezo sensors to provide a rich audio signal to drive the virtual string models. This allows the use of natural strumming and plucking gestures, as well as less traditional gestures such as tapping, scraping or stroking the strings, which have musically meaningful results in the resulting audio signal.

STRINGS WITH SAMPLE TRIGGERING: The sample triggering version uses the same synthesis technique but dramatically reduces the amount of achievable variation in input signal. Rather than passing the audio signal directly to the virtual string algorithm, a peak detection algorithm is used to trigger a pre-recorded pluck recording whenever an amplitude peak is reached. This was the impulse used to excite the physical model. The pluck impulse was recorded directly from the piezo audio signal, and so is directly comparable with the audio-rate version, however the timbre and dynamics remain static independent of input gesture.

TOUCH SENSOR WITH SAMPLE TRIGGERING: For the touch sensor, the same pluck impulse is triggered when one of the six string areas is tapped or swiped across.

## 6.4 STUDY DESIGN

The full study was designed to investigate several factors. Our primary goal was to investigate richness of sensing strategy and its influence on perceived instrument quality. The study design also includes a comparison of the physical form of the instruments (the tabletop and guitar-shaped forms, as described in Section 6.3), as well as sensor topologies (the string modules and touch sensor variations). The richness comparison is only concerned with the instruments which featured the 'string' sensor topology in both 'audio-rate' and 'sample-triggering' versions.

### 6.4.1 *Participants*

There were two groups of participants in this study, non-musicians and guitarists. These two groups were chosen so that one would be familiar with the sensorimotor control patterns that were the *expected* and *desired* means of controlling the instrument (techniques related to guitar playing) [21], and the other group would be physically inexperienced with these techniques. I shall discuss the *Expected*, *Sensed*, *Desired* paradigm in more detail in Section 3.5.5. The guitar was chosen as the base instrument in this study as it is a form that is charged with cultural significance and unarguably a cultural icon due to its centrality in popular music over the course of the 20th and 21st Centuries. This ensured that although the non-musicians would be physically inexperienced with the playing techniques required to control the instrument, they would be familiar with the gestural language associated with the instruments[1].

We recruited 32 participants: 16 guitarists who self-identified as 'competent or better', and 16 non-musicians. 19 participants were male (13 guitarists and 6 non-musicians), and 13 were female (3 guitarists and 10 non-musicians). Participant age ranged from 18 to 62 with an average age of 32. Participants were asked to self-identify at the recruitment stage using the following statements: *'you are comfortable strumming and playing along to a tune'* (competent guitarists) and *'you have no or very little experience playing an instrument'* (non-musicians). In order to account for within-group variability in musical skill, we asked participants to complete the self-report questionnaire of the Goldsmiths Musical Sophistication Index (GoldMSI) test battery [256], the results of which are presented in Figure 6.5. The average GoldMSI scores for each group are shown in Figure 6.5, and were 89 (SD = 11, minimum = 72) for guitarists and 55 (SD = 11, maximum = 70) for non-musicians. The minimum and maximum show the proximity of the two groups. This experiment met ethics standards according to my university's ethics board.

1 Godøy, Haga, and Jensenius [118] have investigated the playing of 'air instruments' and suggest that the imitation of sound-producing gestures by novices can be a rich area of investigation for musicology as it displays a form of *motormimetic sketching* that could be relevant for understanding music cognition.

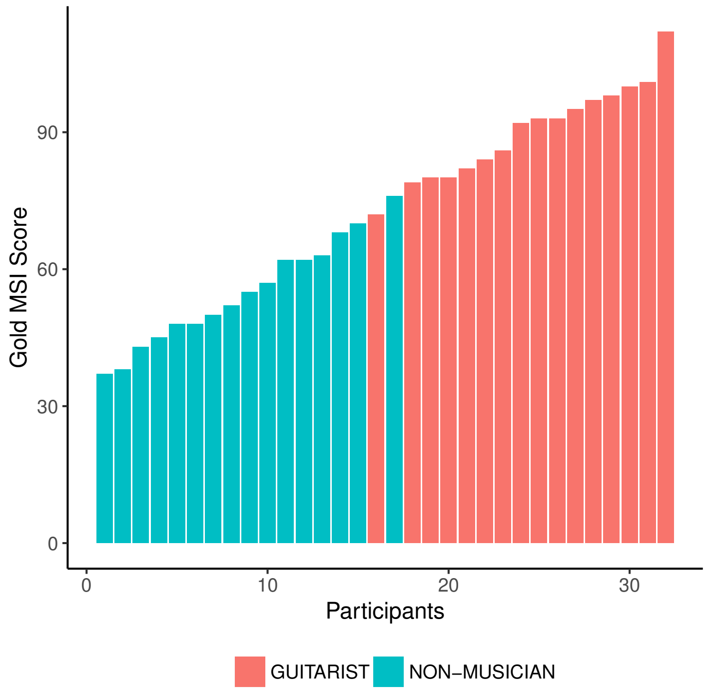



Figure 6.5: The Goldsmiths Musical Sophistication Index test battery score [256] for the participants in this experiment.

### 6.4.2 *Method*

Participants were given one of two combinations of instruments, either the congruent pair SG-TT (Strings-Guitar and Touch-Tabletop) or the incongruent pair ST-TG (Strings-Tabletop and Touch-Guitar). An equal number of guitarists and non-musicians were given each combination, resulting in four groups under test: Guitarists with SG-TT (group A), Guitarists with ST-TG (group B), Non-Musicians with SG-TT (group C) and Non-Musicians with ST-TG (group D). Within each group, the order of presentation of the two instruments was reversed for half the participants.

### 6.4.3 *Experimental set-up*

The study was conducted in the control room of the recording studios at Queen Mary University of London, a sound-proofed space that is usually used for recording music or rehearsals. The sound of the instruments was relayed through a set of active studio monitors that were facing the table where the participants were sitting. None of the participants had prior knowledge of the instruments. They were paid £10 for taking part in the experiment.

### 6.4.4 *Procedure*

The study began with a general introduction via an information sheet that welcomed the participant and advised them that they would be testing some new string instruments and using them to complete some basic musical tasks. The participants were then presented with the first instrument without seeing the second. They were asked to improvise and explore with the instrument and were given 8 minutes alone in the room to do so.

#### 6.4.4.1 *Musical tasks*

Participants were then given a further 8 minutes to rehearse and perform an accompaniment to a recording of a folk song performed on fiddle and electric bass. We chose folk music for this study due to the important role that fretted string instruments play in the genre, as a strummed rhythmic accompaniment, allowing for a relatively accessible musical task to be set up. The choice of a folk tune also provided a strong cultural context to the task. A piece taken from the folk-RNN songbook, a collection of folk songs created using recurrent neural networks [347], was recorded for this purpose. We consulted with an experienced folk musician to choose a song that was stylistically coherent, but that would be unfamiliar to all the participants due to the fact that it was generated using machine learning techniques. The recording had added percussion to make it as easy as possible for participants to follow the beat.

The chord structure of the song used chords I, IV and V in the key of G. We added coloured stickers to the buttons to indicate these chords and printed a colour-coded score for participants to follow while playing. We also produced a video file displaying the chord colours and positions on screen as they appeared in the score, in a similar manner to Guitar Hero games[2]. Participants were allowed to use either or both of these methods to follow the backing track but were encouraged to use the printed score if they felt comfortable doing so. The buttons and score are presented in Figure 6.6. Both the improvisation and score-following tasks were repeated with the second instrument.

For the final musical task, we instructed participants to switch to the audio-rate variation of the instrument using a switch on the instrument's enclosure. They were then given ten minutes to improvise and explore with the instrument. No further score-following tasks were given.

2 https://www.guitarhero.com/uk/en/

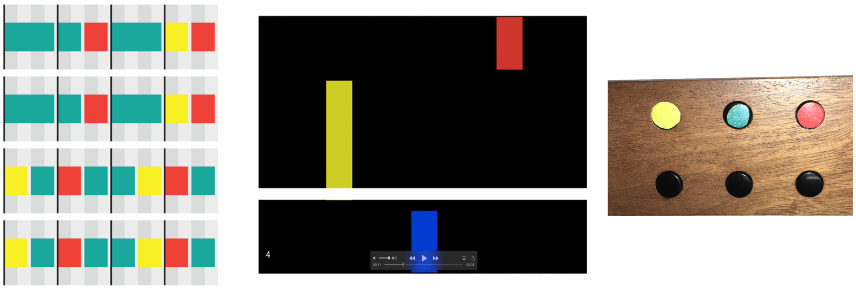


Figure 6.6: From left to right: colour-coded paper score, screenshot of on-screen chord visualiser, colour-coded buttons

#### 6.4.4.2 *Structured interview and questionnaire*

Following the musical tasks, we asked participants to fill out an on-screen questionnaire, providing ratings for each instrument on factors split into technical, social, and general preference subgroups. The questionnaire and results are presented in Section 6.5.4. Following the questionnaire, we conducted structured interviews with questions relating to techniques used with each instrument, and general thoughts on their experiences with the instrument. Following the final musical task with the audio-rate variation, another structured interview took place, this time focusing specifically on differences and similarities between the two variable levels of richness. We then asked participants to indicate their preference for either setting using a horizontal on-screen slider with 'setting 1' (sample-triggering) on the left and 'setting 2' (audio-rate) on the right. This type of CCR produced a value from 0-100, with 0 indicating strong preference for setting 1, and 100 indicating strong preference for setting 2.

### 6.4.5 *Data collection*

Video and audio was recorded throughout the experiment. The video recordings focused on close-ups of participants' hands whilst playing the instrument. Ratings were collected via an online survey on an accompanying laptop that also served as a means of playing back the video score for the musical accompaniment.

## 6.5 FINDINGS

The findings from this experiment consist of a series of quantitative measures from the comparative ratings of each instrument pair and from the comparative ratings of the variable richness settings. The qualitative data gathered during this experiment consists of the two structured interviews conducted with each participant (the first on the comparative rating of the instrument pair and the second on the

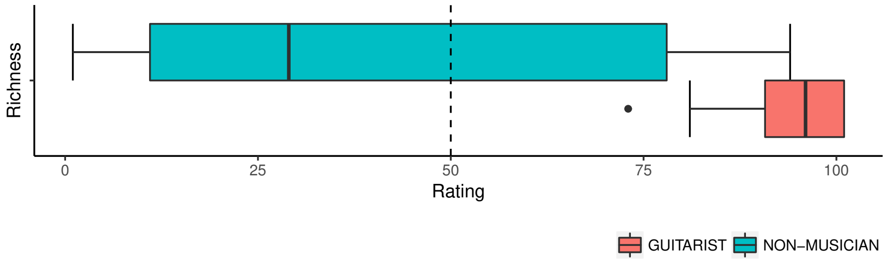



Figure 6.7: Median and IQR ratings of setting 1 (sample-triggering represented by 0 on the y-axis) and setting 2 (audio-rate represented by 100 on the y-axis) for all 32 participants.

variable richness settings). I also present a further analysis of the gestural language of each participant while playing the instruments in order to gauge the difference between participants in terms of gesture usage and the development of gesture over the session.

### 6.5.1 *Richness ratings*

The boxplot in Figure 6.7 shows the comparative rating of the two variable richness settings of the string instrument by each group. A paired t-test on the comparative ratings of the settings from each group found a significant difference between groups ($t = 5.6833$, $df = 16.731$, $p < .01$). All 16 guitarists rated the audio-rate setting as better, whereas there was more disagreement in the non-musician group, with 6 of them rating audio-rate as better and 10 rating sample-triggering as better. What can be seen in Figure 6.7 is a combination of both physical forms, evenly spread between the two groups. A paired t-test of the ratings for each of the two physical forms showed no significant effect of physical form on the results, indicating that the difference in ratings was driven by the difference in richness setting.

A further level of analysis that falls beyond the scope of what is presented in this thesis would be to test for correlations between GoldMSI score and richness rating for each participant. From the GoldMSI scores, presented in 6.5, a continuum of musical sophistication across our sample can be seen, suggesting that a further division of our sample into smaller groups based on this score could result in different trends between them. Given that the focus of this study was the comparison of proficient guitarists and non-musicians I feel the current groupings, with the limit of them being based on self-report of musical capability, suffice for the purposes of this chapter.

### 6.5.2 *Reasoning in relation to control richness*

A thematic analysis was performed on the transcripts from the structured interviews, which focused on the two variable richness settings on the string instruments. This analysis concentrated on the reasoning that people used when talking about the two settings in relation to the following themes: sound, technique, instrument behaviour, relation to existing instruments or interfaces. Table 6.1 presents some sample quotes that are representative of the reasoning of each group.

6 of the guitarists were quick to critique the sample-triggering variation at the end of their session with it, even without the knowledge that a richer mapping would be introduced later in the study. Most of the comments from this group focused on the lack of timbral expression and the flatness of the articulation on the instrument in comparison to the traditional guitar they were used to. The introduction of these capabilities with the audio-rate setting were mentioned by 12 participants in this group. The audio-rate setting's ability to support existing technique was also a reoccurring theme, with particular reference to finger-picking and again to articulation. 6 members of this group also made reference to the 'feel' of the instrument as more 'guitarry' or 'natural' in comparison to sample-triggering: "*setting 2 really uses your knowledge of guitar. Compared to setting 1 where the strings are not behaving as strings*".

The non-musicians who preferred the sample-triggering setting generally gave reasons related to the 'clarity' and 'power' of this setting in comparison to the more 'fragile' or 'far away' audio-rate setting. The sample-triggering setting was described by 9 in this group as easier to generate sound with: the relative force required to create sound with both the instruments was mentioned, with the sample-triggering setting commended for its ability to create a loud sound with little effort and through the use of a diverse set of playing techniques. The audio-rate setting, on the other hand, was referred to as 'difficult', 'hard', or requiring 'too much pressure' in order to produce satisfactory sound. There were also 4 non-musicians (1 who preferred sample-triggering, 3 who preferred audio-rate) who stated that they noticed very little or no difference between the two settings and as a result just followed their instinct.

### 6.5.3 *Playing technique observations*

A further thematic analysis was performed on the video footage captured of each performer. In this analysis I focused on identifying the different sets of gestures each participant used in their right hand during the periods of free-improvisation with each setting. I have conducted this analysis on the three sections of free-improvisation

| TIMBRE | AMPLITUDE | TECHNIQUE | REALISM | BEHAVIOUR |
|---|---|---|---|---|
| Non-musicians | | | | |
| *"Setting 1 was like listening to a guitar in a concert, but setting 2 was more like listening to something on my computer"* | *"Setting 1 was louder and brighter to me"* | *"You need more pressure from your hands with setting 2, it's harder to generate the sounds, with setting 1 you can just touch the strings and create the sound"* | *"With the first setting everything I did made a difference, with the second one everything sounded more fragile in a way."* | *"My tapping didn't trigger the string rather, what I was triggering was the sound of the string itself"* |
| *"You can generate more sound with setting 1"* | *"At first I didn't realise that on setting 1 the volume was only at one level no matter how hard I strummed"* | *"Setting 1 was nice to play because you don't have to pay attention to how hard you strum"* | *"I found for setting 2 that all six strings were playing at the same time"* | *"I did more sliding on the strings with setting 2, but it wasn't easy to get a good sound"* |
| Guitarists | | | | |
| *"There was a definite tonal difference"* | *"With setting 2 you can play soft, and you can play hard"* | *"All the things that I do worked on setting 2 but didn't necessarily work on setting 1"* | *"It really uses your knowledge of guitar. Compared to setting 1 where the strings are not behaving like strings"* | *"Setting 2 was responding to touch much more delicately"* |
| *"Setting 2 sounds like nylon strings from a classical guitar, you can also hear the body."* | *"I have dynamics in general with setting 2, that's a huge benefit"* | *"my gestures were very similar for both settings"* | *"Setting 2 is more like a natural way of playing, with less compression"* | *"Setting 2 is better at picking up picks, it's more responsive"* |

Table 6.1: Selected quotes from the structured interview where participants were asked to justify their ratings of the two variable richness settings.

from each participant: with the touch sensor instrument, the strings sample-triggering setting, and strings audio-rate setting. The observations are presented in Table 6.2. My interest was in comparing the diversity of gesture usage in each group to identify correlations with their given preference. Here the physical form of the instrument has been discounted as a factor to make the comparison of the different input modalities and richness levels clearer; the reported impact of physical form on gesture shall be discussed in a later section.

To begin we shall focus on the difference between the two string instruments with variable richness. From Table 6.2 we can see that there was no clear distinction between the groups in terms of overall diversity of gestures used under these conditions: both seem to use a similar variety and spread of gestures.

guitarists: strings with variable richness As perhaps would be expected, the guitarists used more specialised guitar techniques (strumming as if holding a plectrum, finger-style, palm muting) than the non-musicians. For the guitarists it was observed that between the sample-triggering and audio-rate settings there was in fact a reduction in the occurrence of more exploratory gestures (for example slow plucking or tapping of single strings). This was aside from for the core gestures associated with guitar playing (strumming with hand like holding plectrum and finger-picking). It was also noted that in the case of the guitarists they would generally begin each session with their standard set of guitar techniques and then diversify if they found a certain technique to not be working well, whereas non-musicians displayed no set order in which they introduced different gestures.

non-musicians: strings with variable richness For the non-musicians an increase was observed in the number of participants employing certain gestures that related to the timbre of the instrument with the audio-rate setting; these included scratching the string and muting it while plucking. This could have been due to a number of the non-musicians not knowing what techniques to employ with the richer instrument, reporting it as malfunctioning or drastically reduced in volume. These kinds of gestures align with the tentative and exploratory gestures that were common from this group, particularly slow plucking of individual strings with a single finger.

both groups: touch sensor sample-triggering With regards to the touch sensor variations of the instrument, a broadly comparable diversity of gesture usage between the groups was observed. Guitarists employed slightly more sophisticated gestures that have their equivalent in guitar technique, for example tapping the sensor

| | | Non-musicians | | | Guitarists | | |
|---|---|---|---|---|---|---|---|
| | Instrument: | Touch sample | String sample | String audio | Touch sample | String sample | String audio |
| Strumming | Strumming with a single finger | — | 14 | 12 | — | 9 | 4 |
| | Strumming with hand like holding plectrum | — | 7 | 6 | — | 13 | 14 |
| | Strumming with multiple fingers | — | 5 | 5 | — | 2 | 2 |
| | Rake/'flamenco style' strum | — | 2 | 1 | — | 1 | 4 |
| | Swiping/sliding finger across sensor | 15 | — | — | 15 | — | — |
| Plucking | Slow exploratory plucking with fingers/thumb | — | 16 | 12 | — | 11 | 1 |
| | Finger-picking (finger-style) | — | 8 | 8 | — | 13 | 11 |
| | Plucking with hand like holding plectrum | — | 1 | 3 | — | 2 | 0 |
| | Tapping sensor with single finger | 13 | — | — | 14 | — | — |
| | Tapping sensor with multiple fingers | 4 | — | — | 14 | — | — |
| Scratching/Tapping | Tapping individual strings | — | 10 | 6 | — | 8 | 3 |
| | Tapping multiple strings with flat finger | — | 6 | 2 | — | 4 | 3 |
| | Pushing down on strings | — | 6 | 1 | — | 2 | 0 |
| | Scratching/swiping strings | — | 2 | 8 | — | 4 | 3 |
| | Tapping bridge pieces | — | 1 | 2 | — | 6 | 4 |
| | Palm mute | — | 1 | 2 | — | 7 | 9 |
| | Flat finger tapping across whole sensor | 0 | — | — | 5 | — | — |
| Testing | Damping strings while resonating | — | 6 | 5 | — | 4 | 2 |
| | Strum/pluck/tap at different points along the string | — | 6 | 3 | — | 4 | 4 |
| | Observably testing triggering threshold | 10 | 3 | — | 13 | 4 | — |
| | Observably testing dynamic range | — | — | 2 | — | — | 3 |

Table 6.2: Analysis of gesture occurrence for each group and each setting. Column corresponds to number of participants who used each gesture under each setting. The gesture analysis is of the free improvisation with the three variations: the touch sensor with sample-triggering, the strings with sample-triggering and the strings with audio-rate. A (—) in the table represents cases where a gesture was not possible on an instrument.

with multiple fingers in a manner reminiscent of finger-picking. 5 guitarists also explored some of the less obvious affordances, for example by rolling a flat finger across the sensor to trigger multiple notes quickly creating an effect similar to a flamenco finger rake.

### 6.5.4 *Physical factor ratings*

I shall now report the ratings given by each participant group in relation to the congruent and incongruent instrument pairings, and summarise their reasoning in relation to these ratings given in the structured interviews.

Figure 6.8 displays participants' questionnaire responses given after having played both instruments, which indicate their preferences in relation to the pairing they were given. In what follows the text accompanying each of the questions summarises the trends observed in each group (unless otherwise stated, responses were given by placing a continuous slider on a horizontal plane, with 'Instrument 1' on the left and 'Instrument 2' on the right).

- **Which instrument was easier to play?** For the SG-TT pair, guitarists rated the SG as easier to play, while non-musicians varied in their responses. For the TG-ST pair non-musicians rated TG as easier to play, whereas guitarists varied in their responses.
- **Which instrument allowed you to play in the most natural way?** Guitarists strongly rated strings as more natural on average in both pairings, and non-musicians tended to follow this trend but with more disagreement between participants. There was no noticeable effect of global form.
- **Which instrument was most responsive to your style of playing?** Guitarists rated strings more responsive in both pairings, non-musicians varied in their responses.
- **How well did you play the accompaniment on each instrument?** *(Two 5-point likert scale normalised to 0-100 by making 0 indicate performance is better on instrument 1, 100 indicate performance is better on instrument 2 and 50 represent an equal rating.)* Guitarists rated themselves as playing much better with strings in both pairings, non-musicians varied in their responses.
- **Which instrument was most similar to a guitar?** In the case of SG-TT there is strong agreement between groups that SG is more similar. In the case of ST-TG both groups vary in their responses.
- **Which instrument was more fun to play?** There was no agreement between guitarists, whereas non-musicians rated touch sensor as more fun to play in both combinations.

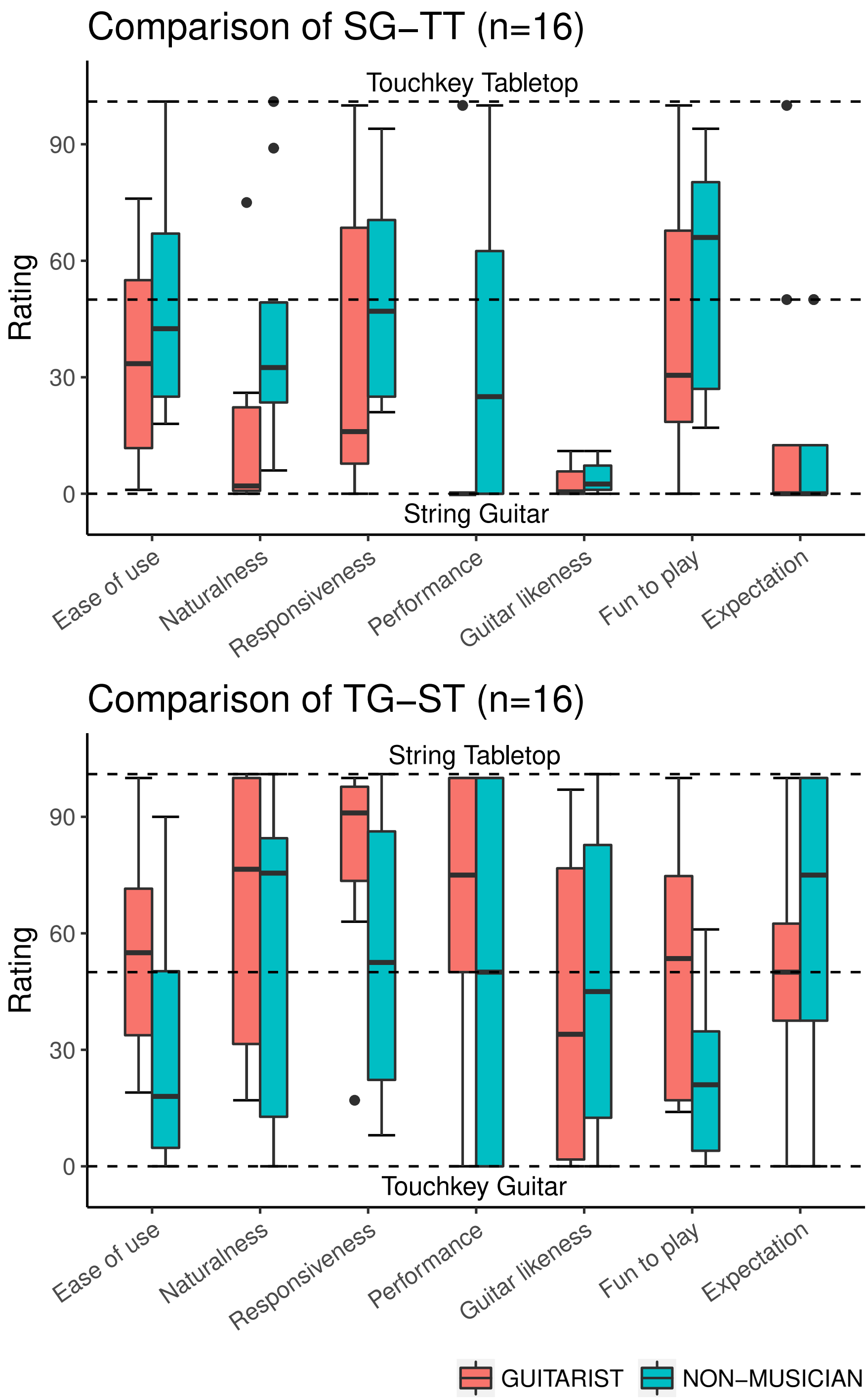



Figure 6.8: The medians and IQRs of the ratings for all participants with each instrument pair.

- **Which instrument best matched your expectation?** In the case of SG-TT there is strong agreement between groups that SG best matches their expectations, in the case of TG-ST both groups vary in their responses.

### 6.5.5 *Reasoning in relation to physical factors*

The structured interviews in which participants gave their reasoning for their ratings of the congruent and incongruent pairings of instruments were transcribed and the responses thematised using thematic analysis. Table 6.3 shows a selection of paraphrased quotes from both groups in relation to the congruent and incongruent pairings, which has been organised into the key themes that emerged from the thematic analysis.

In their comparison of the instrument pairings, 8 of the guitarists referenced the strings as allowing them to use their existing skill in a manner that the touch sensor did not. The lack of tactile feedback on the touch sensor was also a common criticism from this group, with 5 of the guitarists stating that they became lost at a certain point when playing the TG instrument: this was due to not being able to see their hands and having fewer physical anchor points, and was not an issue for this group on the ST instrument. The lack of physical strings made trying to play the instrument like a guitar frustrating, with 3 guitarists mentioning the difficulty of tapping instead of plucking, and the way they had to shape their hand. 2 of the guitarists who had the TG-ST pairing stated that the physical form of the guitar instrument was more important to them than the strings, as it felt strange to have both hands so close together in the case of ST.

8 of the non-musicians also commented on the importance of the strings for their enjoyment of the instrument. They mentioned that by seeing the strings they immediately knew what the instrument would do and how to control it, whereas the touch sensor remained more mysterious and took longer to understand. For others in this group comfort was a reoccurring theme: 6 mentioned that the touch sensor took less effort to play and was easy to control, and 3 made comparisons to touchscreen technology. The importance of the global form of the instrument was mentioned by 10 in this group, who said that when they saw the guitar shaped body, they knew how they were to play. A subset of this group who were presented with the ST instrument first had difficulty understanding how they should play the instrument.

| aesthetics | comfort/ease | comparison | gestures | familiarity |
|---|---|---|---|---|
| | | Non-Musicians | | |
| *'ST looks too traditional, not innovative, but TG looks better'* | *'TG is more comfortable to play but the sensor is not very clear'* | *'ST is kind of like a guitar, it has strings so this looks and sounds like a guitar'* | *'SG seemed much more complicated than TT, there was too much to concentrate on'* | *'With ST I see the strings and know what to do, but with TG it was weird'* |
| *'I know how to hold the SG from looking at it, but not sure how to play the sensor on TT'* | *'I can slide on the touch sensor like a phone screen, it's easy* | *'I would play TG on stage, ST seems much more experimental'* | *'I mostly pressed the strings with ST, only realised I should strum after playing TG'* | *'TG felt more like holding a guitar, I could just strum away'* |
| | | Guitarists | | |
| *'TG said guitar to me as soon as I saw it. With ST I thought of lap steel'* | *'TG is more comfortable as a guitar player because of the position of the arms'* | *'ST is much more preferable because it's like an autoharp, importantly it's got the physical strings'* | *'Everything was more direct with ST, with TG I kept losing my position'* | *'The tactility of the strings [ST] lends itself to more innate skills'* |
| *'TG felt more like a game controller, but ST felt like a real instrument'* | *'The resistance of the strings on SG was important for strumming'* | *'Whereas TT was fun to explore, as a guitarist I preferred SG'* | *'I tried to play fingerstyle with TG but tapping in time was more difficult than plucking ST'* | *'The position of the bits of the instrument in TG was important, but once I adjusted to ST the strings made the difference'* |

Table 6.3: Selected paraphrased quotes from the structured interviews where participants were asked to compare either the congruent or incongruent pairing of instruments.

## 6.6 DISCUSSION

In the discussion that follows I shall first address the question of variable richness, followed by the influence of the input modality and physical form in the instrument pairings.

### 6.6.1 *Control intimacy*

The findings from this experiment complicate the notion that 'a richer instrument is a better instrument'. The guitarists in this experiment were unanimous in their preference for the richer setting, which is perhaps unsurprising as the audio-rate setting more accurately translates the existing techniques of guitar players to musically meaningful timbral and dynamic effects. What was less expected was the ambiguity in the responses from the non-musicians, and their overall tendency to prefer the less rich instrument.

From the structured interviews it is possible to piece together a picture of why this difference in opinion might exist: the guitarists were able to speak at length of a lack of detail and a flattening of nuance in the less rich setting, and pointed out what they regarded as the general shortcomings of this setting before they even knew a richer version was coming. For this group there was a wasted reserve of gestural potential that had no effect on the musical output of the instrument. This resulted in a degree of frustration as the physical supports of the instrument implied a rich set of gestures that the instrument, in its sample-triggering form, was unable to handle.

Many of the non-musicians, on the other hand, were complimentary of the sample-triggering setting for its clarity and strength of sound. The simplicity and variety of gestures that could be used to create a clear plucking sound (tapping the strings or bridge pieces, plucking in a variety of ways) with the instrument was treated as a positive aspect of its design. When this group encountered the richer audio-rate settings there were frequent reports of the quietness of the instrument (which was true for soft playing, but if sufficient energy was put into the instrument it could be louder than the sample-triggering), of a lack of brilliance in the sound of the instrument (due to the fact that timbre was now under their control), and of the instrument becoming more delicate, and harder to produce a satisfactory sound with.

### 6.6.2 *Gestural behaviour*

The similar spread in gestures between both groups, alongside the stark difference in opinion in terms of setting preference, shows how the instruments' constraints were encountered in a similar manner for both groups, but meant different things to them depending on experience. This is a point that we shall explore in more depth in Section 6.6.5.

For the guitarists we observed a general decrease in the diversity of techniques employed in relation to the richer setting, aside from the core gestures associated with guitar playing (strumming with hand like holding plectrum and finger-picking). There are two possible explanations for this reduction. With the richer setting, guitarists had less need to deviate from standard guitar-based techniques, which generally worked well with this setting, and thus continued exploring these gestures rather than using a more diverse palette of gestures. Alternatively this effect could have been due to the order of presentation of the two settings, with the richer setting always coming second, and the fact that the participants had already settled on their favoured techniques. It was also noted that the guitarists would generally begin each session with guitar techniques and then diversify, whereas the non-musicians introduced new gestures in a more random order. This suggests that the lower diversity of gestures of the guitarists with the richer setting could be mostly due to them getting satisfying results from the instrument with the techniques they wanted, or expected, to use.

The non-musicians were more concerned with *efficiency* of sound production, and used any means possible to make the instrument create sound. In the case of the string sample-triggering instrument various types of tapping on the strings and bridge pieces or scratching the strings to produce sound were witnessed, a pallet of gestures that is far removed from the standard set of techniques that would be used with an acoustic guitar. With the audio-rate setting these gestures became less useful, as this setting required the performer to take full control of the sound production and put more energy into the instrument, yet they were still used by this group, albeit with diminishing musical satisfaction.

The touch sensor represents an input modality with much lower expectations from the performer in terms of sensorimotor skill. This kind of sensor is commonly found in touchscreen devices and is intuitive and straight-forward to use in a way that strings are not. Once the active area of the sensor was identified by both groups they had no problems producing sound by strumming the sensor: to do so simply requires laying fingertips on the sensor. Accordingly there was less diversity in the gestures employed by both groups: participants generally swiped or tapped the sensor, after an initial period of testing the string regions that divided the sensor into six. Guitarists were more musically explorative with this sensor, tapping in a manner reminiscent of finger-picking. A common complaint leveraged by the guitarists concerned the lack of tactile feedback with this sensor (withstanding the rubber ridges) which led them to drift away from the active area of the instrument.

### 6.6.3 *Instrument efficiency and learning curve*

If we return to Jordà's notion of instrument efficiency [167] as discussed in Section 3.5, we could say that the sample-triggering setting is more efficient than the audio-rate setting. The complexity of musical *input* is matched for both settings: both have reasonable coverage of the techniques used to play guitar, and the physical layout and dynamic material behaviour of the strings do not change with the settings, both physically supporting the same base of gestures. It is in the musical *output* complexity that the settings differ, with the sample-triggering setting projecting a rich and nuanced set of input gestures to a reduced set of musical features in the sound output. The audio-rate setting retains the spectral relationship between input and output and thus requires nuanced control of the strings in order to get a nuanced output, whereas the sample-triggering setting can work with a much lower level of definition at the input: any energy above a certain threshold is transformed into an impulse and the spectral signature of that gesture is disregarded.

Another angle to view this from is in terms of required *effort*. Both settings transfer energy from input to output (the kinetic energy of the performer to the sound energy of the instrument) and effort relates to the total *amount* of energy needed to achieve a result. This influences how the performer moves with the instrument: Ward [370] proposes that the body movements created by a DMI should not be left to chance, and that in fact designers should consider the physical effort involved on the part of the performer and the physical resistance that an instrument imposes on the movement of the performer. This builds on Vertegaal, Ungvary, and Kieslinger's [365] suggestion that physical effort is an essential component of musical interaction, both for the performing musician and perceiving audience. In the case of this experiment the sample-triggering setting required much less physical effort to achieve an equivalent note than the audio-rate setting. In the case of non-musicians the relative lack of effort was perceived as attractive with the rich setting becoming more difficult. In the case of the guitarists, they reported that something was missing, which we could posit was related to the effort required to perform on the instrument and the resultant nuance of musical control that this grants.

The findings from this experiment highlight the balance between expertise and challenge in an instrument's design. In 3.5.2 we discussed the learnability of DMIs and touched upon Csikzentmihalyi's flow theory [64]. The variation in observations can be explained in terms of this theory. Considering the high challenge of the string based instruments and the low challenge of the touch sensor based instruments versus the respective abilities of each group, we can consider this in terms of balancing skill and challenge. The instrument that is most enjoyed is the one that matches challenge and ability, and maintains flow. This also echoes previous work by Nash and Blackwell where they suggest that the best interfaces are those which are able to scale challenge as ability develops [259].

For the experienced guitarists who are used to navigating an instrument in which complexity is maintained from input to output there is no advantage in reducing it, in fact many of the guitarists negatively commented on the lack of output complexity from the sample-triggering setting. In the case of the non-musicians, there does however seem to be a use in reducing the output complexity (and hence increasing the efficiency of the instrument). Part of the reason could be due to the learning curve that each instrument has. Whereas the guitarists already know a large amount of techniques that can be used to play the guitar, and so are able to quickly make sense of the audio-rate setting, non-musicians want a more direct route to producing musically satisfying sound with the instrument and so prefer the more shallow learning curve of the more efficient sample-triggering. The switch on the instrument that changes settings can be considered as a kind of 'complexity management' [284] that provides non-musicians with a short-cut to taking part in a musical activity as a performer.

### 6.6.4 *Input modality and instrument form*

Results from the questionnaire show the effects of the congruent vs. incongruent pairings as they relate to input modality and physical form. The focus of this thesis is on tangible experience when controlling DMIs and so for that reason the input modality and its effects on the participants are of most relevance, and physical form shall only be discussed briefly. A full exploration of the influence of physical form on the acceptability of these instruments within an existing musical practice can be found in Harrison et al. [130]. In the experimental design the SG-TT pairing represents a congruent relationship: it is clear what the guitar form with strummable strings is designed to do and this was generally reflected in the ratings of both groups and their comments during the interviews. On the other hand, the ST-TG pairing introduces an incongruence between form and modality and this seems to have the effect of dividing the two groups of participants in their opinions of the instruments: the non-musicians rated the TG instrument (which shared the physical form of a guitar but lacked strings) more favourably, whereas the guitarists gave higher ratings to the ST instrument that had physical strings but a compact tabletop form. The differentiation here seems to depend on the input modality, and it is on this that we shall now focus.

The reasons for the different preferences between the two groups are interesting as they highlight the design elements that carry the cultural weight of an instrument. The guitarists generally gave higher ratings to the stringed versions of the instruments regardless of global form, and from the structured interviews there were a number of reports of the strings feeling more natural to play and allowing the use of existing techniques that they knew from the guitar. The tactility of the strings was also mentioned as an important factor, for the way they provided a physical support to gestures. This group's criticisms of the touch sensor repeatedly focused on the lack of an anchor or reference point that would tell them where their hand was positioned, leading their hands to drift away from the sensing area if they were not visually monitoring the instrument.

The non-musicians were much more diverse in their reports of preference: there was a relatively even split between the two input modalities. For the string instruments there was an increase in unconventional techniques observed for this group (tapping on bridge pieces, tapping and pushing down stings, flat rolling of fingers to trigger strings). This group was in fact much more inventive in their interpretation of the strings than the guitarists, who generally stayed close to their learned techniques. The touch sensor input modality was rated as more fun to play than the strings by the non-musicians regardless of global form, suggesting that the novelty of this interaction could have advantages with this group. Although the strings represent a certain canon of musical gesture, this is easily completely side-stepped when the performer has no knowledge of this canon.

### 6.6.5 *The projection*

In what follows the focus is on the effects of the variable input modality and control intimacy, interpreted in terms of the projections between performer and instrument.

input modality Let's begin by considering the two input modalities of the instruments: the touch sensor and strings with vibration sensing. In each case the movement that the instrument receives is by definition different: the position and approximate pressure of a fingertip on the touch sensor; the amount of vibration in the string as registered by the piezo disc. These represent two different projections onto the same output parameters (the physical model of the string), but via a different sensing strategy, physical support and mapping. From the gesture analysis performed in Section 6.5 clear categories of techniques emerging in response to the two different input modalities were noted.

The types of gestures that were used with each instrument varied depending on familiarity with the instrument and existing expertise. Non-musicians, for example, were more prone to use the strings in unconventional ways (tapping on them, scraping or pushing them down to trigger samples). For this group there was no existing coupling of action and sound in relation to guitar technique and the physical form of the strings. Rather than being concerned with the 'fit' of existing gestures on the instrument, they were much more likely to settle on an unconventional technique that was able to produce sound reliably after a period of exploration. The guitarists on the other hand seemed more intent on investigating how their existing practice related to the constraints of the instrument, and in the case of both input modalities participants from this group began their exploration of the instrument by testing the instruments with guitar techniques (as discussed in Section 6.5.3) before diversifying. Similar trends were displayed with the professional percussionists as discussed in Section 4.6.5. However for the touch sensor input modality the behaviour of the two groups ended up following similar paths, perhaps because the affordance structure of the sensor was relatively clear, and well known to both groups from interacting with touchscreen devices. The differences between the two groups highlight an interesting point about design that echoes Benford et al.'s expected, sensed, desired framework [21]. Often gestures that fall outside of the category of the expected and desired are nevertheless chosen by performers due to their efficiency in producing sound.

In terms of the instruments' design it is the haptic feedback (both tactile and kinaesthetic) of the input modalities which supports existing sensorimotor patterns. Both instruments contain a good deal of detailed physical feedback that the performer can perceive through their sense of touch: even the touch sensor, the more tangibly neutral of the two instruments, contained tactile guide points on its surface in the form of raised ridges where each of the strings would be triggered. There is enough tactile feedback to give the performer a good idea of the position of the strings in both cases, although it is more subtle on the touch sensor. The kinaesthetic feedback that the performers get from each of the instruments, i.e. their awareness of how their body moves with the instrument, is also roughly equatable. If we consider the *meso* and *macro* gestures of the performers (as discussed in Section 2.2.3) they generally exist within very similar territory for both input modalities. 'Strumming' is realised through a motion of the hand and arms is that is roughly the same whether swiping across the touch sensor or strumming the strings, and 'fingerpicking' happens through similar movements of the fingers with the hand anchored by the sensing area. It is the *micro* elements of the gestural language that are altered: strumming with the hand held as if holding a plectrum becomes swiping with the fleshy part of a fingertip; finger-picking strings, instead of relying on a plucking action, is achieved by tapping the sensor in the correct places.

In relation to the material presented in Section 2.1.4 on the coupling of action and perception through musical training, it may be possible to hypothesise that in the case of the guitarists in this experiment, they displayed a tight coupling between action and perception that is built around the physical structure of the strings of a guitar. Their hands knew the distances of the strings, how to play them in sequence, and a whole series of articulatory gestures and nuances that relate to the dynamics and timbre of a note as it is excited. With the guitarists and the touch sensor, the physical feedback was shifted and although macro aspects of the playing technique were maintained, guitarists reported being unable to put their existing sensorimotor skill to good use. The relative acceptance of the touch sensor-based instrument amongst the non-musicians may be due to the fact that they lacked the strong coupling of action and perception that revolved around the form of the strings and thus didn't know the *choreographies* associated with this physical arrangement.

VARIABLE RICHNESS An instrument's aperture decides what parts of a performer's action passes through the instrument and what does not. In the variable richness settings of the string instrument we could describe these variations as either a widening or narrowing of the aperture. The two variations were either passing the full audio signal captured by the vibration sensors to excite a physical model of a string, or using this same signal to trigger a sample that was used as excitation. The physical characteristics of the instruments remain identical but the manner in which the captured movement is mapped to the sound engine changes, changing the projection of the instrument.

In terms of maintaining existing techniques, I found that the richness of the transferral of action to sound was an important factor for the guitarists. Although the physical supports of the strings gave the impression that guitar techniques would be able to be employed in the control of the instrument, the sample-triggering version of the strings was quickly deemed as inappropriate for a whole set of guitar gestures by the guitarists that related to nuanced articulation and timbral variation.

Returning to the concept of experiential control [355] presented in Section 3.5.5 it is possible to describe this as a shifting of *push* and *pull* effects. For the guitarists a whole set of pull effects are introduced that allow them to employ a gestural language that they know well and that gains musical meaning with the wider aperture and richer transferral of action to sound. With the sample-triggering setting the guitarists seemed to find push effects in the simplifications in the instrument's design in comparison to the instruments that were used to: the narrower aperture was interpreted as a lack of bandwidth in their musical control. For the non-musicians we see the opposite happening in many cases. With the sample-triggering setting, they were quickly drawn to the pull effects of the instrument – the places where they were enabled in their control [355]. The additional richness of the audio-rate settings acts like a barrier to them being able to create meaningful music on the instrument: the extra effort and control that the richer setting necessitates is outside their familiar sensorimotor experience, and so the additional richness acts more like a push effect, disabling them in their control of the instrument.

The aperture of the audio-rate instrument can be imagined as wider than the sample-triggering version: more of the performers' gestural language can be projected through the instrument. A wider aperture might inherently mean lower *efficiency* [167] (as discussed in Section 6.6.3): the amount of input and output complexity are better balanced, and the performer, in their sensorimotor control, becomes responsible for the extra effort required to produce musically meaningful sound from the instrument.

In DMI design we have the great benefit of being able to adaptively widen or narrow the aperture of an instrument. Situations in which the aperture of the instrument matches the skill level of the performer could be seen as similar to attempting to stay in the flow state [64], the area of interaction where challenge and skill are balanced allowing a performer to avoid both anxiety and boredom. Another question that this experiment raises is whether and how a projection changes with the player and their skill. The musical potential that a performer brings to an instrument is dependent on their training and this greatly influences the projections that are possible. When a performer has extensive musical experience they are bringing more to the instrument in the first place, they have a richer source of body movements for the instrument to project downwards.

## 6.7 CHAPTER CONCLUSIONS

This chapter has explored the notion of 'control intimacy' in an instrument's design, and its relation to instrumental expertise and to prior sensorimotor experience. In this case control intimacy has been interpreted as the richness of an instrument's sensing strategy, and a comparison case was made between two string instruments: one with sample-triggering based on amplitude and another that used a full audio signal to drive physical modelling synthesis. Findings and observations from the experiment presented in this chapter support the notion that for experienced musicians richer instruments are preferable and deemed more suitable for performance. From the reasoning of the participants, this seems to be due to the preservation of the full spectrum of gestures that have a meaningful effect on the musical output.

When designing for non-musicians, however, the role of richness is less clear: there was a greater spread in overall preference for the two instrument variations, with a tendency towards the less rich version, and instances in which in the difference between the two settings were not noticed. I have discussed how the less rich variation of the instrument can be considered as the more musically *efficient* of the two settings, especially in the hands of a non-musician, and these findings suggest that an instrument requiring less physical and mental effort from the player can lead to more enjoyable experiences for beginners, even if it makes available a more restricted space of possibilities. This is perhaps because the less rich instrument grants inexperienced performers streamlined access to a sophisticated form of music-making that is beyond their current level of sensorimotor skill. The value of learners taking part in more informal, less technique-oriented musical situations has been recognised [241].

This experiment has analysed the 'first contact' of a performer with an instrument rather than a longitudinal evolution of the performer-instrument relationship. In Section 3.5 I spoke of the notion of 'complexity management' [284] in instrumental learning. We may find a compelling argument for introducing similar methods through the adjustment of the richness of future instruments, progressively introducing more richness as the performer gains sensorimotor skill. Future studies on this subject could therefore incorporate a longitudinal approach and a sliding scale of control intimacy.

The influence of input modality and physical form on judgements of an instrument's quality was a secondary question that this experiment addressed, and perhaps as expected the findings in relation to this are a little less clear. This could be due to our attempt at separating global form and input modality, resulting in a series of instruments where the divisions are not clean-cut: each instrument in a pair has an influence on and echoes the design cues of the other. In terms of input modality we can conclude that the physical presence of a set of strings is important for guitarists, who displayed a preference for the technical familiarity of the strings, even when they were not behaving like they normally would. For non-musicians there is less allegiance to a particular input modality, with their preferences shifting more towards the cultural load of the guitar form regardless of input modality.

From this experiment it is possible to consider control intimacy as a tangible factor. A number of instances in which varying the richness of an instrument created effects that can related to the tangible experience of the instrument were observed: physical aspects of an instrument's design can change in their meaning to a performer depending on the richness of sensing. Physical supports are important for encouraging patterns of pre-existing musical control, whether they are known in a sensorimotor sense or simply in a cultural sense. When it comes to musicians, what becomes crucial is how these supports are treated and activated as part of the instrument's *virtual mechanism*. Richness can be considered as a similar issue to action-sound latency, but while latency relates to the temporal coupling and the tightness of the bond between action and sound, richness relates to bandwidth, the amount of nuance of sensorimotor control that passes through the instrument to the sound output.

# Part III

# CRITICAL REFLECTIONS

# 7 THE PROJECTION MODEL

*This chapter incorporates material from 'Rich gesture, reduced control: the influence of constrained mappings on performance technique' by Jack, Stockman and McPherson, originally published in the proceedings the 4th International Conference on Movement Computing, MOCO 2017 [157].*

One of the primary aims of this thesis is to characterise the sense of touch during interaction with digital musical instruments. This involves clarifying the perceptual bases that are brought together as 'tangible experience' when performing with an instrument, and identifying the design parameters that influence and modulate this experience, as outlined in Part i. In Part ii we looked at three practical studies that each investigated a particular manifestation of tangible experience when performing with a DMI. Each of the studies was based around a probe DMI and analysed the encounter of differently skilled groups of performers with the instrument. In particular the studies analysed impressions of instrument quality in relation to tangible design elements, and examined performance parameters relevant to the tangible cues under investigation.

This chapter serves to further reflect on the projection model as presented in Section 3.6. It presents a post-hoc analysis of the instrument from Chapter 5, in which the projection model is used as an analytical and discursive tool to reflect on the interaction of performer and instrument in this case. The focus will be on the implications of the tangible design cues in this instrument for the patterns of action that emerge around the instrument.

The projection model is a conceptual tool that offers forward language and terminology for describing what is happening in the coupling of performer and instrument. The model has its basis in *embodied control* (as introduced in Section 2.4.3), that is, how a performer uses movements in the control of an instrument, and reciprocally, how the movements of a performer are influenced and even controlled by the instrument [355]. As with the other models we discussed in Section 3.5, this model is not designed to incorporate all of the levels of detail that each individual model can highlight in their specific focus. Rather, this model is introduced as a complimentary perspective on the design of DMIs, that hopes to guide others building similar instruments and provide a vocabulary for discussing tangible elements of DMI design.

## 7.1 PERFORMER-INSTRUMENT COUPLING

A musical instrument, whether digital or acoustic, only allows a small fraction of the total 'information' available in a performer's actions to pass to the sound it creates. In Section 3.2.3.1 I discussed how musical instruments can be considered as *movement transducers*: the 'active surface' of an instrument translates a certain subset of a musician's movement into the behaviour of sound waves [15]. Every instrument converts the actions of a performer into sound in different ways: the design of acoustic musical instruments is led by an exploration of how the acoustic properties of certain materials can be put at the service of sound production, whether through striking a naturally sonorous object in the case of idiophones, striking a membrane coupled to a resonant chamber in the case of membranophones, creating standing waves in a tube in the case of aerophones, exciting vibrations in a string that are then amplified and filtered by a resonant chamber in chordophones [188]. Acoustic instruments balance the sensorimotor constraints of a performer's movements with the acoustic constraints of the materials. In DMIs on the other hand, nothing comes 'for free', and almost everything must be accounted for in design [215]. In each case certain categories of actions become sound-producing while others have no effect on the sounding behaviour of the instrument.

From an information-theory point of view we could say that an instrument represents a communication channel that allows a certain bandwidth of information to pass through, and has a set of constraints that define which parts of the movement that it receives 'bear meaning'. In other words, the information that an instrument receives as an input is the movements of the performer on the active surface of the instrument, and those that are judged as meaning-bearing from the perspective of the instrument are the ones that result in an audible effect. In the case of the DMI the communication channel of the instrument consists of a combination of the physical structure, sensing strategy, mapping strategy and sound generation of the instrument. In the descriptive model presented in Section 3.6 instead of using the language of information-theory a metaphor of projection is used.

This projection is enabled by a lens whose characteristics decide on how action passes to sound: which actions are brought into focus, magnified, blurred or refracted, as well as the action space that is captured in the first instance through the lens's orientation. As the actions of a performer are projected through the instrument they pass through a point of reduced or minimum representation before being projected back outwards as the sonic and kinematic behaviour of the instrument. The instrument's behaviour contains a signature, at least to some degree, of the initial action responsible for causing it.

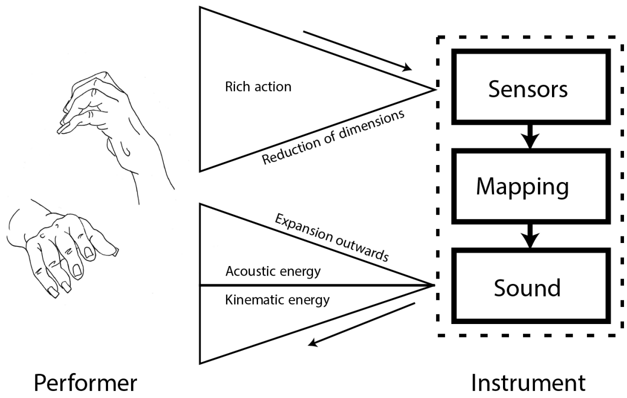



Figure 7.1: A representation of the projection model.

## 7.2 PROJECTING THROUGH AN INSTRUMENT

In Section 3.6 the three elements of the projection model were introduced: *projection* as the objective physical possibilities between a body and an instrument, *aperture* as an objective descriptor of an instrument and the flow of action through it, and *choreographies* as the patterns of movement that form around an instrument.

### 7.2.1 *Projection*

In this model performer-instrument interaction can be understood as a kind of projection from a higher dimensional space to a lower dimensional space. A projection is defined as a property of a performer-instrument coupling, and represents the objective space of possibilities between a body and a sound, grounded in the physical instrument. As can be seen in Figure 7.1, the input to this model is the actions of the performer which contain many biomechanical degrees of freedom and many event-based decisions made by the performer about how they interact and how this relates to the wider environment.

The projection between performer and instrument is as if caused by a lens: parts of the action of the performer are passed through transparently, parts are refracted and distorted, other parts are brought into sharp focus or become blurred and out of focus, while others do not pass through the lens at all. The manner in which this projection happens – the reduction in dimensions and characteristics of the lens – are direct results of an instrument's design and the fit of a performer's body to that instrument. A projection represents the elements of a performer's actions (whether sound-producing or not) which are maintained as musically meaningful and which are rendered meaningless.

In the transformation from a 3 dimensional to a 2 dimensional space, for example, there is an axis along which changes would not be observable in the output. In a similar manner we can imagine factors such as the orientation of the lens to the actions of the performer influencing the types of action that are meaningful in a musical interaction. Actions may be brought into sharp focus or become blurred and out of focus and this relates to ideas of control intimacy discussed in Section 3.3: the degree to which the actions of the performer can pass through the lens of an instrument undistorted and the clarity with which they can do so.

### 7.2.2 *Aperture*

The aperture determines the overall flow of light that is able to pass through an optical model, and so in this case it can be understood as determining the bandwidth of action that can pass through the instrument into sound. In the design of DMIs the resolution with which action is registered and passed through the instrument can often be quantified due to the sampling rate of the sensors and the way this is aligned with the audio. In Section 3.4.3 we discussed the different rates at play during an interaction; by considering the aperture we can open up conversations about the flow of energy in an instrument, and about how it is sampled and quantised. It is possible to imagine the aperture opening or closing depending on the amount of action that is allowed to pass through a projection. The width of the aperture influences aspects related to transparency and control intimacy, central topics of Section 3.3. This point is discussed further in Section 8.2.

DMI designers finely tune the lens and aperture responsible for the projection between performer and instrument. An aim of this model is to promote a deeper consideration of these factors, particularly the aperture which represents the minimal flow of action to sound in the instrument. This minimal representation of action is a fundamental characteristic of all DMIs that heavily impacts on an instrument's character.

### 7.2.3 *Choreography and idiomaticity*

*Choreography* is taken to mean an emergent pattern of interaction which results from the performer's subjective engagement with the projection and aperture of the instrument. The choreographies that emerge over the course of an interaction with an instrument represent its idiomatic gestural space for a certain performer. This is to do with the ease of playing certain passages, the performance practices that fit an instrument, and how they emerge over a period of time in relation to an instrument.

The term *choreographies* is taken from Loke and Kocaballi and Tuuri, Parviainen, and Pirhonen's notion of *technology-induced choreographies* [204, 355], a term they use to describe how movements are used in the control of technology, and how reciprocally, movements are controlled by technology. Technologies, from their perspective, contain *pre-choreographic* inscriptions, predefined patterns of action that are then enacted through interaction when the choreography is performed. As discussed in Section 3.6 the colloquial understanding of the term has pros and cons and a more neutral term such as *kinetographies* could be a good alternative with a focus on movement and without the associated connotations of dance. That said, we find it more useful to align the intended understanding of this term with that of Tuuri, Parviainen, and Pirhonen and so have remained with this terminology.

De Souza describes the formation of idiomatic actions as follows: "[t]he idiom is realized in players' overlearned actions, in the ways they typically move through an instrumental space, revealing some affordances and concealing others" [68, p. 77]. His definition builds upon the earlier exploration of idiomatic organisation in music from Huron and Berec: "[w]hat makes something idiomatic is not that it is easy to play, but that it is easier to play given the specific prescribed circumstances compared with other possible performance circumstances" [149, p. 119]. Idiomaticity, then, has to do with the way action projects through the instrument, and the relative ease or difficulty with which a action or sequence of actions produce a sound. Performers adjust their actions to project through an instrument, settling into idiomatic territory based on the fit and the comfort of a particular action with the mechanism of an instrument, and its musical meaning.

## 7.3 POST-HOC ANALYSIS OF PERCUSSION INSTRUMENT

This section aims to apply the projection model as a conceptual and descriptive tool to a specific interaction context: the group of professional percussionists playing the digital percussion instrument from Chapter 5. I present a post-hoc observational analysis of the professional percussionists' actions in relation to the instrument and their development over the course of a one hour session. This instrument has a projection that can be considered a classic constraint of DMI design: velocity triggering. By focusing on the variation and evolution of musical gesture I aim to identify the design parameters that influence this projection, demonstrating how the gestural language that the performers use is richer than what the sensors ultimately capture, and analysing how they adapt their movements to project them through the instrument as a choreography begins to establish.

### 7.3.1 *Digital percussion instrument*

The design of the instrument used in this study is described in detail in Section 5.3. In this instrument there is a layer of electronic decoupling, where sound is a direct result of making contact with the surface of the ceramic tiles. The electronic decoupling serves the same purpose as the mechanical decoupling of the piano, which is to say that the interaction between the hand and the surface is extremely rich and gives rise to quite a wide variety of sounds and effects, and yet in the case of the tiles it is projected down by way of a single vibration sensor and a peak detection algorithm.

PROCEDURE This section presents an analysis of the first and last 15 minutes of the hour-long encounter that each of the professional percussionists had with this instrument. The majority of the encounters consisted of the percussionists being asked to focus on the difference between two randomised settings on the instrument in several different performative contexts (free improvisation, rhythmic tasks, structured improvisation). The settings which involved variable amounts of action-sound latency did not encourage the percussionists to choose any macro set of gestures over another, aside from an impact on subtle details of their temporal performance (discussed in Chapter 5). Both the first and last 15 minute sections involved free-improvisation.

DATA COLLECTION AND INTERVIEWS The entirety of each session was audio and video recorded. The session ended with a 15 to 20 minute structured interview where the percussionists were asked to describe their experience of the instrument, the techniques used throughout the session and their impressions of what worked well and what did not.

### 7.3.2 *Performance analysis*

The video recordings of these sections were analysed using thematic analysis [40]. Figure 7.2 shows an example of the manual annotation of the video recordings that was performed, and the subsequent thematisation of gestures. The first pass of annotation involved identifying the descriptive categories that best explained the range of sound-producing gestures used. This included *motor effector* (the part of the hand that made contact with the instrument plus the parts of the body that were in motion as part of the ancillary gesture), *temporal be-*

*haviour* (the distinction between individual strikes and compound gestures like rolls, multiple finger strikes and fast finger-thumb or finger-finger combinations), *local spatial behaviour* (the location of strike on an individual tile), and *global spatial behaviour* (the arrangement of the hands on the instrument as a whole). Each of these gestural characteristics were identified on the time-line of the video and coded with a relevant tag.

A similar thematic analysis was then performed on the interview material which allowed me to reinforce the codings from the video and identify established techniques and styles of playing. When taken in isolation the observations of gesture from the video analysis do not necessarily give the full picture of the decision making processes the percussionists employed. Accordingly, after having reviewed the interview material, I performed a second pass of annotation, this time focusing on the stylistic techniques employed and sound-facilitating gestures. The stylistic categories then were built from the techniques and instruments mentioned during the interviews and from my personal knowledge of percussion performance.

### 7.3.3 *Observations: the evolution of gesture*

In Table 7.1 I present a summary of my observations from the thematic analysis of the performances.

motor effector From the motor effector category in Table 7.1 we can see a summary of the ways in which contact was made with the instrument. By comparing the fourth and fifth column it is possible to identify the effectors that persisted throughout the session and those that did not appear in the final 15 minutes. Many techniques that have minimal acoustic impact on the output of the instrument, but that would be meaningful techniques on an acoustic percussion instrument, stopped being used: scratching, playing with fingernails, dampening of tiles, drumming on the rim of tiles with the sides of straight fingers, playing with the palm, swiping the finger across the tile like a touch screen. The techniques that continued to be used were those with a clear and direct sound-producing function. This includes playing with the thumb, multiple simultaneous fingers, finger-thumb combinations and, most commonly, playing with fingertips.

local spatial exploration All 10 percussionists explored the local response of at least one tile in the first 15 minutes and only 1 percussionist in the final 15 minutes. Local spatial exploration was about testing the active areas of the tiles and finding the physical constraints of the instrument's design: repeated strikes were typically performed whilst moving from the centre to the rim of the tile. As

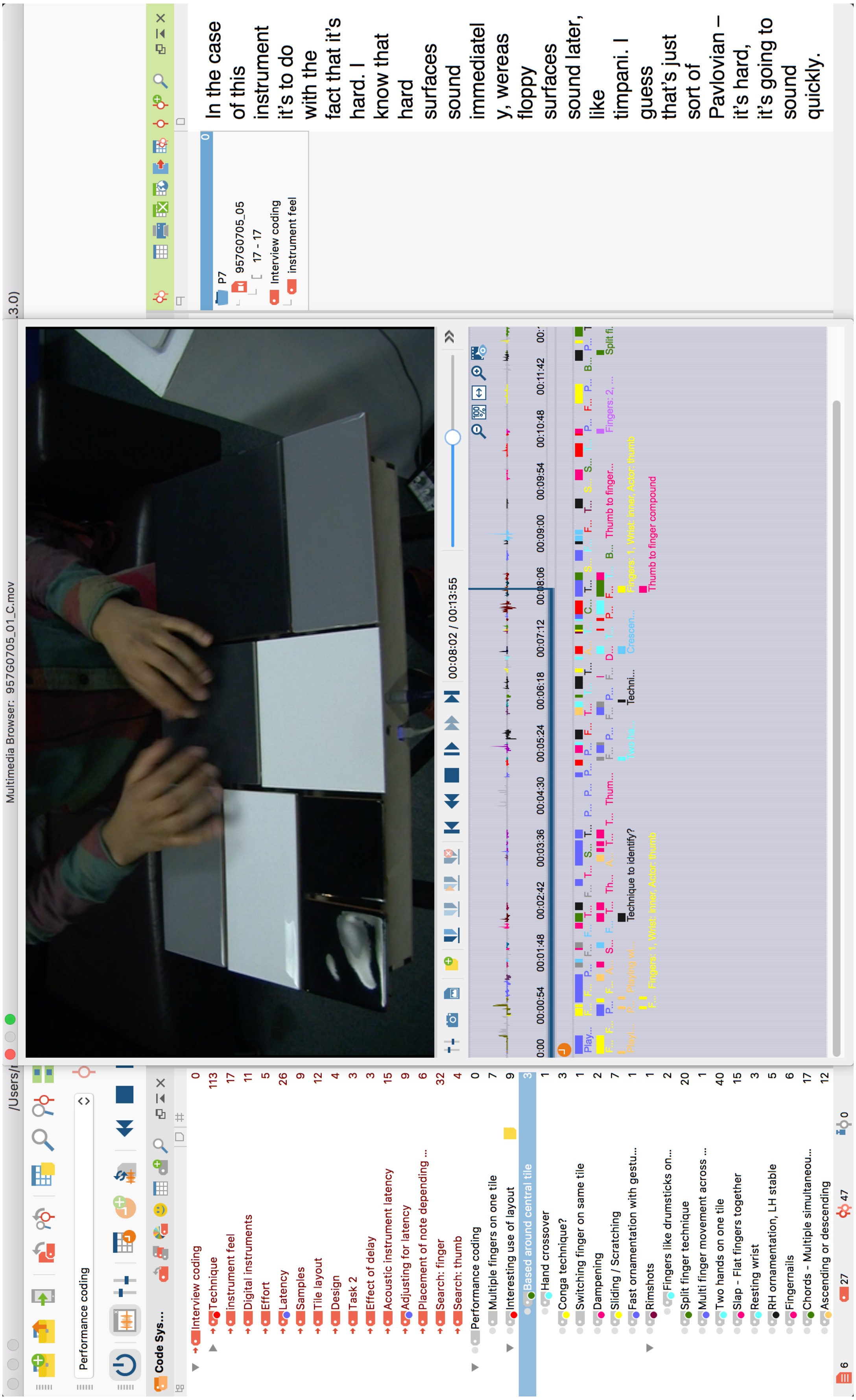



Figure 7.2: An example of the annotation process of the video recordings of the percussionists. Themes and coding structures can be seen on the left.

| | Description | Acoustic effect | Start | End |
|---|---|---|---|---|
| Motor Effector | fingertips | ✓ | 10 | 10 |
| | fingernails | — | 9 | 1 |
| | thumb | ✓ | 7 | 4 |
| | palm | — | 3 | 0 |
| | dampening | × | 6 | 0 |
| | drumming on rim with side of fingers | ✓ | 2 | 0 |
| | flat hand slap | — | 6 | 2 |
| | scratching tile | × | 4 | 0 |
| | touch screen swipe | × | 2 | 0 |
| Compound gestures | finger-thumb combination | ✓ | 4 | 4 |
| | chords | ✓ | 6 | 4 |
| Temporal | single hand rolls on one tile (fingertips) | ✓ | 10 | 3 |
| | two hand roll on one tile | ✓ | 10 | 3 |
| | strumming fingers (straight) | ✓ | 6 | 0 |
| | split hand technique | ✓ | 6 | 3 |
| Spatial | exploring tile surface area | × | 10 | 1 |
| | hand anchored on central tile | n/a | 2 | 2 |
| | splitting the interface in thirds | n/a | 2 | 2 |
| | ascending / descending patterns | n/a | 2 | 1 |
| | free movement across interface, no anchor point | n/a | 10 | 10 |

Table 7.1: Analysis of gesture occurrence. The column titled *start* shows the number of percussionists who used this technique more than once in the first 15 minutes, *end* does the same for the last 15 minutes. Techniques marked with a dash produce a sound from the instrument but there is no difference in the quality of that sound compared to the more commonly used motor effectors. The symbol × represents the case where the technique had no acoustic effect.

the response did not vary much locally this kind of gesture did not provide useful acoustic differences in the signal. The scale of the tile was also a factor in the local exploration patterns of the percussionists. In the interviews 3 percussionists mentioned that for bigger tiles they would have expected a wider variety of sonic responses, for example different active areas or varying response toward the edge of the tile, which was not the case. When asked, all 3 stated that this was not an issue for the smaller tiles.

global spatial exploration The global spatial arrangement of the percussionists' movements was one of the important indicators of the established percussion technique that they were employing. The geometrical layout of the tiles encouraged the percussionists to arrange their hands in different ways in relation to the instrument, including splitting the interface into sections with different musical purposes (down beats on one third (left hand) and ornamentation on the other two thirds (right hand)) or anchoring their movement around a central tile so that they were always moving away from the same position. The most common pattern of global spatial movement was free movement across the interface with no anchor point – this was employed by all percussionists in both the first and last sections. The greater reduction in variation of local in comparison global exploration implies that the percussionists came to realise that there was not a lot of musical function in local spatial variation, yet varying their global behaviour gave them the ability to play in established percussion styles.

established technique A summary of the main established percussion techniques that were used is presented in Table 7.2. For a percussion technique to be successfully supported by the instrument its base components had to have an acoustic effect, and the macro movement patterns of the style had to be supported by the layout and physical characteristics of the instrument. Techniques carried into the final 15 minutes and named by the percussionists in the interview were as follows: tarabuka split hand technique (sharp strikes with fingertips, single hand rolls with fingertips of curved fingers, wrists rotating), hang drum playing technique (detached strikes with fingertips, middle or index finger), tar (frame drum) technique (Middle Eastern split hand technique, again all fingertips with straight finger rolls, no wrist rotation), conga thumb-finger technique (compound gesture with thumb-finger combinations[1]).

Techniques that did not have an acoustic effect and were subsequently removed from the gestural vocabulary of the performers are as follows: tabla strokes (relies heavily on palm muting, resting wrists

1 Note that palm strike and flat-handed slap were removed from this style of playing in most cases – see the next section for further discussion.

| | Description | Acoustic effect | Start | End |
|---|---|---|---|---|
| Stylistic technique / comparable instrument | Tabla strokes | × | 3 | 0 |
| | Cuban/Latin finger style | × | 4 | 0 |
| | Djembe | — | 3 | 0 |
| | Tarabuka | ✓ | 7 | 3 |
| | Hangdrum | ✓ | 7 | 5 |
| | Tar | ✓ | 1 | 1 |
| | Conga | — | 3 | 2 |

Table 7.2: Stylistic techniques and comparable instruments from the second pass of the video analysis and interviews. The dash symbol represents the case where some gestures work but are timbrally flattened. The symbol × represents the case where elements of the technique had no acoustic effect.

and varying note duration with dampening from the striking finger); Cuban/Latin finger percussion (relies on heel-finger motion, rocking the hand back and forth with combinations of palm and flat finger slaps); djembe techniques (relies on palm hits, slaps and changing the tension of the drum with a dampening finger).

### 7.3.4 *Interviews with percussionists*

The percussionists were asked to describe the techniques they found more or less effective on the instrument. The interview was conducted in front of the instrument with demonstrations encouraged. They were also asked how the instrument related to other instruments they had played (acoustic, electronic and digital) and to established hand percussion techniques.

Two distinct categories of reasoning for the success or failure of certain percussion techniques became clear from my coding of the interviews: *acoustic* and *ergonomic*. In Table 7.3 I present some paraphrased examples of this reasoning. The acoustic reasons tended to address the instrument's ability to produce an accurate response to a style of playing, and so included comments about the sensing and sound engine. For example: "*I noticed with the fingertips it's more accurate and easier to get feedback*", "*when you hit it you know what volume you'll get*", "*there was something more musical about using the fingers rather than the palm*", "*to get the most reliable impact and the most control I used my index finger*".

Ergonomic reasoning tended to address the physical form of the instrument, its layout, the ceramic as a playing surface, its stability or durability. For example: "*this sort of motion is really nice and natu-*

*ral*", "*when using hands it's better to have something harder to dig into*", "*I was doing a lot of fast stuff with these two* [demonstrates top two right tiles] *and then longer notes here* [demonstrates the left two thirds of the instrument]", "*I tried a little bit of thumb action but I found it a bit too clunky*". The percussionists also related the tile instrument to other percussion instruments and established percussion techniques with which they had experience. For example: "*you can do all the tarabuka stuff, also Middle Eastern tar* (frame drum) *playing*", "*it's similar to clay udu drums in the hardness of the tiles*", "*I can play it like a hang drum, all fingertips*".

### 7.3.5 *Discussion*

In this section I will frame the observational analysis presented above in the terms outlined in the projection model.

#### 7.3.5.1 *The acoustic and the digital*

With the exception of the swiping gesture derived from touch-screen interaction, all of the sound-producing gestures came from acoustic hand percussion instruments. Hand percussion has a large but nevertheless limited vocabulary of gestures that involve many different parts of the hand and various striking patterns – a set of techniques that are predicated on a richer transfer of action to sound than this instrument affords. When these choreographies, which were developed with other instruments, meet this instrument which is more constrained than most acoustic instruments, a reduction in the variety of gestures employed was observed. The remaining sets of actions are aligned with the constraints of the instrument. After interacting with the instrument for a while certain actions are decided to be more musically useful than others, and the performers' action are reduced to a smaller set of techniques related to the behaviour of the instrument. This reduction involves the percussionists employing a subset of the degrees of freedom that they would expect from an acoustic instrument, suggesting a reduction in the bandwidth of interaction. We can describe this as the percussionists adjusting their actions so they can project through the instrument into sound.

The instrument, instead of creating its own choreography from scratch, initially borrows patterns of action from the instruments that it is most similar to. This is in part due to the high degree of training of the percussionists and their specialism in percussion, but also is due to ecological factors regarding familiarity with everyday interactions [76] and enactive aspects of how action translates to sound [80] as discussed in Section 3.4. This instrument conforms to physical principles about the transferral of energy and errs towards a simula-

| Reasons for ceasing to use a technique | | Reasons for continuing to use a technique | |
|---|---|---|---|
| *Acoustic / Sensors* | *Ergonomic / Physical* | *Acoustic / Sensors* | *Ergonomic / Physical* |
| The instrument didn't respond well to this type of technique | This type of gesture was uncomfortable to play / didn't fit | When I play like this the sound is just right there | The feel of the tiles was just right for this kind of playing |
| I couldn't hear the sound that corresponded to this technique | The size/layout of the tiles didn't allow me to play like this | I liked the range of dynamics that this technique allowed | The layout suggested that I arrange my hands like this |
| This technique didn't change the sound | This gesture was too forceful for this instrument | Like this I was able to play quickly and feel in control | The tile size encouraged me to play with this technique |

Table 7.3: Summary of paraphrased reasons for why percussionists continued to use or discontinued a particular technique.

tion of an acoustic instrument but differs in some fundamental ways: limited timbral variation, each tile remaining a trigger, limited active areas, no possibility to dampen notes. Yet from the comparison of this instrument with an acoustic percussion instrument we could posit that the projections in both cases are similar: the possible couplings between action and sound are shared to a large degree between the two instruments. It is in the characteristics of the lens responsible for the projection that the instruments differ, that is in how actions are carried into sound, the transparency of the lens, the focus drawn on certain actions and blurring of others, and the aperture that decides on the overall flow of action.

Time is another factor in the development of a choreography. Within the span of an hour the performers go from techniques derived from traditional percussion instruments to a narrower and more focused set of techniques that seem catered towards the specific characteristics of this instrument. Over a longer period of time, based on previous work by Gurevich, Stapleton, and Marquez-Borbon [125] and Zappi and McPherson [389], we might expect to see performers coming up with novel techniques that take advantage of non-obvious affordances specific to this instrument and techniques optimised to draw the most from the main affordances of this instrument, an appropriation that didn't have time to form in this case.

#### 7.3.5.2 *Voluntary self-constraint*

From my observations and from the percussionists' reports it is possible to make some hypotheses about the experiential control of the performers: their subjective experience of the instrument which I shall discuss in terms of the *push* and *pull effects* [355] as discussed in Section 3.5.5.

PERFORMERS DISCARD ACTIONS WHICH HAVE NO MEANING ON THE INSTRUMENT (PUSH EFFECTS) Dampening the surface of the tile was a technique employed multiple times by over half the percussionists. The instrument's physical construction and initial behaviour led the percussionists to believe that it afforded this action. However, the projection of the instrument did not pass information related to this gesture: the action had no sonic effect and so held little musical meaning. As a result percussion techniques that required dampening as an essential element also stopped being used by the percussionists. Scratching the surface of the instrument is another example where the percussionists seemed to expect some kind of timbral change, an affordance that the instrument did not have and hence scratching was not used by the end of the session. These are both examples of push effects in the instrument's design: the performer's expectations of how action should relate to sound output are disrupted, pushing them towards an alternative set of performance techniques.

PERFORMERS CHOOSE THE MOST ERGONOMICALLY CONVENIENT WAYS TO PLAY (PULL EFFECTS) The pull effects of the instrument are displayed in the most widely employed control strategies across all of the percussionists – playing with the fingertips and performing fast single-hand rolls on a single tile, for example. The percussionists identified these techniques as the most convenient and effective ways of playing the instrument, where they were certain about the quality of output they would receive from their input gesture and could trust how the instrument was tracking their gestures.

PERFORMERS OPTIMISE THEIR ACTIONS TO BE MAXIMALLY EFFICIENT WITH RESPECT TO THE INSTRUMENT'S CAPABILITIES (ALSO PULL EFFECTS) Alongside the influence of physical constraints on what is needed to actually play an instrument, we can also consider there being a further constraint that is self-imposed by the performer on the basis of what they find to be musically meaningful or not while playing. In this case the percussionists' initial strategies were borrowed from many different percussion techniques, each of which is best adapted to a different instrument. These choreographies are carried onto the tile instrument and adapted or optimised to be best suited to the projection of that performer and instrument.

As an example, a flat-handed slap which is a common component of conga playing was kept within the gestural language of two performers, but not due to it having its expected timbral effects (a sharp and dampened strike). Instead it was used to accentuate notes in amplitude. As it fitted within the global established percussion practice of conga playing, for which most of the other elements worked well on the instrument, it was kept within their choreography but changed in musical purpose. Performers may choose certain strategies on the basis of what's ergonomic and convenient and in relation to the training they have, but they will also choose their strategy on the basis of what they believe is musically meaningful in relation to their coupling with the instrument.

We could describe what is happening here as the performers reflecting on the projection that is formed from the coupling of their body and the instrument. The projection reflects back onto the actions that the performer uses and they self-constrain their actions based on this response. As a choreography forms this is based on the actions brought into focus, magnified, blurred and occluded by the lens responsible for the projection.

### 7.3.6 *Projection in the case of this instrument*

In Figure 7.3 I illustrate the specific projection of this instrument as an inverted pyramid that projects downwards from a space of rich body language to the aperture of the instrument and then outwards to sound and haptic output. On the right hand side of the illustration I have noted some of the aspects of performer action that are discarded at each step of this projection.

In order to trace this specific projection it is best to work upwards from the aperture, which includes the measurement of vibration on the tile and its reduction to discrete velocity triggering. The vibration on the tile is approximately equivalent to the strength of an impact between that tile and another object – timbral qualities are rejected from the mapping. There are many different parts of the body that can make contact with the tiles and trigger notes, many of them with the same or very similar effects as read by the sensors: information again is lost here. For any given part of the body that strikes the tile there is a complex kinematic system going up the fingers, hand and arm which could produce an identical type of collision: this is another place where information is discarded. From the perspective of the aperture all motion of the performer's body above the plane of the instrument is irrelevant, except insofar as it is needed as a preparation to strike the tiles.

When designing digital musical instruments the designer takes charge of the aperture of the instrument and the lens responsible for projecting from action to sound. Recognising that the aperture is an inherent feature of an instrument's design can help designers to decide how they deal with it and how, through both active and passive elements of the instrument's design, they can maintain the rich action of the performer. Even within the traditional paradigm of sensor-based DMIs, without attempting to quantify what gesture means, considering the instrument as the locus of a projection downwards can influence design decisions: thinking from the top of the projection (the space of potential action-sound couplings grounded in the instrument), to the final musical behaviour may help make decisions about the kind and quality of musical action instrument builders want to maintain through the design of an instrument.

#### 7.3.6.1 *Physicality and material concerns*

While I have compared the instrument used in Chapter 5 with acoustic percussion instruments, this study could have been equivalently done with any of numerous commercially available drum pad interfaces. With a small rubberised plastic drum pad certain types of gestures (single hand rolls in which the fingers are strummed across the tile surface, or split hand percussion tarabuka technique) would not be supported. I can posit that the ceramic interface, vibration sensing strategy and fast response of the system all contributed to the maintenance of these particular gestures. The difference lies in the warping of the energy transfer curve, which is inherent in piezo discs rather than force sensitive resistors (commonly used in drum pads), in the geometric disposition of the instrument which changes the gestural language in comparison to a grid of equally sized squares, and in the use of ceramic for both the feel and expectations as opposed to rubberised plastic.

In Section 3.6 I spoke of the importance of the aperture, as responsible for controlling the overall flow of action. Identifying the point of minimal flow of action within an instrument's design is an important descriptor of an instrument which helps to clarify the passage of action to sound characterise the instrument. Every instrument has a point of minimal flow and in the case of DMIs this point can often be clearly identified. In the comparison of the tile instrument to a rubberised drum pad interface, the aperture can be said to be more or less the same size (in that they are both velocity triggering), but the projection is different on account of the different materials and geometry, and hence different choreographies form.

The choreographies of the percussionists were observed as becoming more focused and tailored towards the affordances of the instrument but nonetheless remaining much richer than the sensors can actually capture. Instead of through the sensors, much of the diversity in the choreographies that form with this instrument are implicitly encouraged by its physical design. Physical design factors are central to the projections observed. The way these factors are reinforced through mappings create pull effects on the instrument: they support the performers in their control of the instrument, enabling them to employ their tacit knowledge of such an interaction and their prior training [80] and acting as metaphors of control present in the instrument's material [143]. Layout and size are also crucial factors: they restrict and guide the trajectories of movements, the speed with which it is possible to move across the instrument, the patterns of triggering that are possible. These aspects of the instrument's design go beyond obvious mapping descriptors yet are crucially important on the basis of the percussionists' actions that were observed.

## 7.4 further projections

As demonstrated in the previous section the projection model can lend useful terms and concepts for describing the meeting point of performer action and physical design. To further explore the applicability of the model I shall now discuss some different instrumental scenarios in these terms.

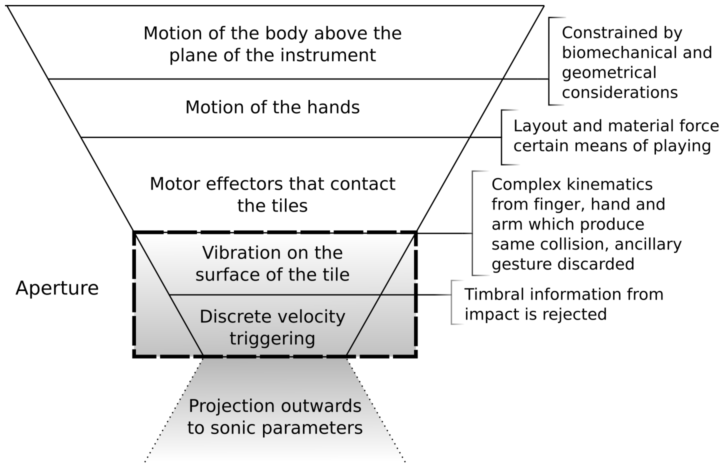



Figure 7.3: A representation of the specific projection in the case of the ceramic tile instrument. On the right-hand side are the elements discarded at each layer of reduction. The aperture of the instrument is highlighted.

### 7.4.1 *Augmentation and appropriation*

THE PIANO Let us return to consider the situation of the piano. The piano has a layer of mechanical decoupling between the control actions of the player and the sound production. The richness of touch could be measured in many ways, for example as continuous key angle [234], but in the case of an acoustic piano is projected by the escapement mechanism, and by the lever into ballistic motion. In this case we have an extremely rich action, as discussed in Section 3.6 in relation to the work of Doğantan-Dack, which remains rich right up to the level of the physical contact between the player and instrument. From that point onwards gesture becomes discretised to note-onset and velocity, projected down to a set of sonic features and behaviours. Due to the aperture of the instrument actions lose meaning, for example using multiple fingers on one key, any finger movement on the key surface once the hammer has been activated, any type of plucking, scratching or strumming the key surface that might be used on an instrument with a more direct coupling of action and sound. Static aspects of a piano's design are responsible for maintaining a great deal of the stylistic gestural language associated with the instrument. For example the asymmetrical arrangement of the keys provides a terrain that the performer can navigate with anchor points at every octave, equally the octave is approximately scaled to the reach of one hand.

THE MAGNETIC RESONATOR PIANO Let us now consider the case of instrumental augmentation. The magnetic resonator piano (MRP) is an electronic augmentation of the grand piano created by Andrew McPherson which uses electromagnets to induce vibrations in the strings independently of the hammers. This makes a series of new performance techniques available: infinite sustain, crescendos, harmonics, pitch bends, all controlled from sensing at the piano keyboard [235].

In the case of the MRP the aperture of this instrument is opened up in comparison to that of the traditional acoustic piano. As augmented instruments often aim to leave existing playing techniques intact, without obstructing or interfering with them, we could imagine that the general characteristics of the projection of the traditional piano remain very similar to the MRP's projection. However with the MRP the keyboard sensing and actuation could be seen as opening up the aperture, creating more opportunities for musical gesture to translate to musical sound: this augmentation allows more and different actions from the performer to be projected through the instrument. The aperture has been opened up for certain types of movement that wouldn't pass through the traditional piano, for example slow depression of the keys without engaging the hammers. This augmentation also slightly alters the overall characteristics of the projection in respect to the piano, in that some techniques that previously made sense on the piano no longer make sense on the MRP, for example the keys and pedals producing different types of sustain.

INSTRUMENTAL CHANGE AND APPROPRIATION In the case of instrumental change, for example extending the grand piano through the inclusion of screws, marbles and bolts on the strings, the instrument itself will be changed and the possible projections through the instrument altered. The performer here is intervening as luthier. As an example let us consider the case of Jimi Hendrix playing electric guitar with feedback (for an ecological reading of Hendrix's performance see [59]). The first question to answer is about the boundaries of the instrument in this case: we could say that the whole system of electric guitar, guitar effects and amplifier are what constitutes Hendrix's instrument rather than any of these part considered individually.

When performing with this system which is highly dependent on distortion, the volume of the amplifier and the feedback that occurs between the electric guitar and amplifier, a projection is still at play. In comparison with say, more traditional jazz electric guitar playing, a different type of signal, one sustained via feedback, can project through the instrument which would not have been possible with a more traditional set up. In this sense we can speak of the aperture opening with the addition of the effects and amplification. On the other hand, the projection has grown to encompass all of these technologies in a different way than before and this enables new performer-instrument choreographies to form.

### 7.4.2 *Style and idiomaticity*

violin versus fiddle Let us now consider the difference between the western classical violin and Scottish traditional fiddle, two styles of playing, or choreographies, that can be accommodated on the same physical violin. The projection of a performer and instrument describes the space of objective possibilities between a body and an instrument. Scottish traditional fiddle, Hindustani classical violin and western classical violin each have distinct and well-defined performance practices on the same physical instrument. In each case the holding of the instrument and bow, correct posture, and sitting position differ and project action through the instrument to sound in different ways.

In this respect the projection through the instrument is subtly different in each case: after extensive training in a particular performance style the performer has formed strong action-sound pathways that change the potential held in their actions in comparison to someone who has not received that training. Idiomatic styles of playing regard the fit between a performer's body, the physical characteristics of the instrument and the musical style they are playing in. This is heavily dependent on the training a performer has been through and the musical context that they have developed within. With time and training choreographies of action form around an instrument.

different schools of piano playing In a similar manner we can imagine how different schools of piano playing train their pianists to move their bodies with great precision in ways specific to an institution or even to a maestro. These gestural languages represent different projections through the piano, which contribute to different choreographies that are formed as the performer learns to inhabit the musical space of the piano. Piano pedagogy is full of examples of performance techniques that represent subtly different projections to the same sonic output as we discussed in Section 3.6.

### 7.4.3 *Applicability of the model*

In this chapter and in the discussion sections of the chapters in Part ii the projection model has been used as a descriptive tool for discussing how instrument design and performer action relate, and how idiomatic styles of playing are established on an instrument. The projection model tries to find appropriate language to describe DMIs and the way they encourage patterns of action. The aim of this model is to act an an interpretative tool which can draw attention to the choices made by designers, particularly design decisions relating to physicality and tangibility, and to demonstrate how performer-instrument coupling can be described as a projection that must pass through a point of reduced representation. This chapter has applied the model to the post-hoc observation analysis of a particular performer-instrument interaction.

## 7.5 SUMMARY

This chapter has introduced a model of performer-instrument interaction based on projection: an additional perspective on the exchange that happens between performer and instrument, one which aims to foreground embodied control [355]. The model proposes that all musical instruments, acoustic, electronic and digital, involve a projection of sorts: the rich action of the performer meets the constraints of the instrument which projects that action down to a reduced representation via the aperture which decides on the flow of information through the instrument. This is then re-expanded out to the rich sound and kinematic output of the instrument. The designed affordances and constraints of the instrument behave like a bottleneck in the interaction, in that they allow specific actions of a musician to pass through and be translated into the musical response of the instrument, and prevent others from doing so.

The projection model provides a way of describing and analysing the action of a performer as they engage with an instrument, and aims to help detail how an instrument selects for a certain cut and quality of a performer's actions. This model works towards giving tangible and physical design parameters a more central place in the analysis and conceptualisation of DMIs and to highlighting issues relating to physical behaviour, craft and subtle tangible design differences between DMIs.

# 8 CONCLUSIONS AND FUTURE WORK

## 8.1 SUMMARY OF RESEARCH

This thesis has shed light on some key aspects of physical interaction with DMIs and on the role that rich manual manipulation plays in this exchange. It has explored instances in which the sense of touch is integral to musical interaction, using control intimacy – the degree to which the sensorimotor capabilities of a performer and the behaviour of an instrument are coupled – as a vehicle for discussing the finer details of *tangible experience* during musical control. Defining 'tangibility' is a work-in-progress across many different fields of research, and the work presented in this thesis has studied the relationship and interpretation of this concept through the analysis of performance on a series of probe DMIs. These have been put to the test with performers with different levels of training, identifying a number of issues that influence the tangibility of a DMI, or its *feel*.

The aim of this research has been to lay out foundations for considering the design and analysis of DMIs from a primarily tangible perspective. To this end, this thesis proposes a series of recommendations for designing DMIs able to create a rich physical experience for performers. A series of practical design parameters and implementation techniques that influence this experience have been identified, each of which have been discussed in the context of current research on musical sensorimotor control and skill acquisition. This research also considers tangible experience across different levels of musical skill, presenting findings that relate to expertise and sensorimotor familiarity. This involves appreciating the nuance and sophistication of movements that seem effortless, like pushing a button, pressing a key, turning a dial, striking a surface, and trying to account for this richness in design by maintaining the touch of a performer throughout the system of a DMI.

Each of the individual experiments presented in this thesis contain insights into a number of central aspects of the design of new DMIs. In particular they highlight the fundamentally interdisciplinary nature of the design of such instruments, which unite research and techniques from HCI, perception studies, musicology, engineering and experimental psychology. This research, too, has taken a multidisciplinary stance and for this reason has adopted a methodology that treats DMIs as technology probes, which has allowed access to 'hard-to-reach' areas of interaction. This approach has enabled an analysis of performer-instrument interaction that functions on two different levels simultaneously: the empirical level of performance studies and systematic musicology, and a more evaluative level that relates to embodied interactions within HCI and design more widely. This thesis has also questioned the importance and distinctness of musical interaction when considered within the bigger picture of interactions between humans and computers. Part of the reason why DMIs make such a good test bed for HCI generally has to do with the tight coupling that exists between action and sound that musicians develop over many years, and the highly skilled and complex example of tool use that playing a musical instrument represents.

In Section 8.2 of this chapter I summarise the main findings and recommendations from the experiments in this research. Section 8.3 outlines the main contributions of the research, Section 8.4 reflects on the methodology that has been followed in this research, and Section 8.5 suggests some directions for future work.

## 8.2 DESIGNING FOR RICH PHYSICAL EXPERIENCE

The work presented in this thesis contains a series of themes and observations on tangible interaction that I hope will inspire and inform future research in this area. Here I present the key elements of tangible design that each of the studies presented in Part ii have highlighted.

### 8.2.1 *Control intimacy*

A central concept of this thesis is that of control intimacy, which was discussed at length in Section 3.3. At various points in this thesis I have argued for its importance when considering tangible interaction with DMIs. The degree of control intimacy that an instrument has can be considered the strength of the coupling between action and sound in an instrument: intimate instruments utilise a large part of the spectrum of a performer's sensorimotor capabilities and put them in service of music-making. The connection of body movement and musical expression and cognition, as discussed in Section 2.2.1, puts forward an argument for strong action-sound couplings in DMI design. By making the nuances of a performer's movement hold more presence in a musical interaction, intimacy represents the amount that a performer can bring to bear through their physical interaction.

#### 8.2.1.1 *Richness of sensing*

In terms of *instrumental control* [355] as discussed in Section 3.5.5, intimacy relates to the resolution with which movements are captured and represented digitally within the core of the instrument. Practically, this is a question of sampling rate and bit depth of sensing, and of how this resolution is maintained at various stages in the digital system of the instrument. It is also a question of the physical layer between the performer and the sensor, which can respond with greater or lesser nuance and stability. In Chapter 6 I presented a series of instruments with varying levels of intimacy via either the reduction of a captured movement into a control signal for triggering a sample, or the use of the auditory signature of a captured movement to directly drive a physical modelling synthesiser. The biomechanical movement of the performer is ostensibly the same in both cases, however the treatment of this movement by the digital system of the instrument differs dramatically. The gestures of the performer on the active strings of the instrument generate an acoustic signal that is treated differently in each case: in the first case all timbral information generated by the control gesture does not pass through the *aperture* of the instrument which treats this movement as a quantity of amplitude alone; in the second case the full spectrum of the auditory signal of the gesture passes through the *aperture* and influences the behaviour of the instrument. In Section 6.6.5 I discussed how these two instruments differ in terms of the width of their apertures, and the implications of this for the movements of different groups of performers.

In terms of *experiential control* [355] this varying richness impacted on impressions of instrument quality in a manner highly dependent on expertise. When comparing two instruments with variable richness I saw a pronounced difference between the guitarists (who unanimously preferred the richer setting), and non-musicians (who generally preferred the less rich setting, and a quarter of them could not tell what was changing). Findings and observations from the experiment presented in Chapter 6 support the notion that for experienced musicians richer instruments are preferable and deemed more suitable for performance: from their reports in a structured interview this seems to be due to the preservation of the full spectrum of gestures that have a meaningful effect on the musical output of the instrument.

When designing for non-musicians, however, the role of richness is less clear. In Section 6.6.3 I discussed how the less-rich variation of the instrument could be considered as the more musically *efficient* of the two settings. Experience as a confounding factor was also discussed in relation to flow theory and the manner in which skill and challenge are balanced in an instrument's design. My findings suggested that an instrument requiring less physical and mental effort from the player can lead to more enjoyable experiences for beginners, even if it makes available a more restricted space of musical possibilities. A challenge of DMI design is striking the right balance between richness of sensing, the efficiency of sound production and the learning curve of the instrument.

Varying sensing richness can make physical aspects of an instrument's design change in their behaviour, modulating the perceived effort needed to play an instrument and changing the instrument's *virtual mechanism* (see Section 3.4). The additional bandwidth of interaction manifested itself differently for the two groups. For the non-musicians in the study presented in Chapter 6, there were reports of the instrument falling quiet, becoming fragile, and performance becoming more difficult with the richer setting, whereas for the guitarists the richer setting was reported as more alive, more responsive, and as responding to touch in a more delicate manner. From these above reports we can identify the tangible behaviour of the instrument modulating between the variable richness settings: the perceived effort required to play the instrument and quality of the instrument change substantially, while the physical arrangement of the instrument remains unchanged.

In his theorisation of tangible user interfaces Horn points out that the evocation of cultural forms can tap into users' existing cognitive, physical and emotional resources, activating existing forms of social activity [143]. In relation to the guitarists observed in this study it is possible to consider control intimacy as part of the 'cultural form' of the guitar: a guitar-like instrument that does not match the level of control intimacy that a trained guitarist is accustomed to does not have the same ability to tap into a performer's existing sensorimotor and affective resources.

#### 8.2.1.2 *Action-sound latency*

Another integral aspect of control intimacy (as previously discussed in Section 3.3) is the temporal determinacy of an instrument's response: a requirement for intimate control is the close temporal coupling of action and sound. As discussed in Section 5.1, this concerns a number of issues relating to the design of a DMI and the manner in which the timing of movement information is dealt with within the instrument's digital system. Action-sound latency has been widely recognised as a barrier to fluent interaction with a DMI, but there is a lack of experimental evidence as to the effects of latency on performer experience, an issue addressed in the study presented in Chapter 5.

Importantly, this experiment has shown that the negative effects of action-sound latency on an instrument's perceived quality do not depend on a performer's ability to consciously perceive a delay between action and sound, nor on the temporal acuity of the performer. While the professional percussionists who took part in the experiment presented in Chapter 5 were significantly more literate in their descriptions of latency, and more perceptive of latency in comparison to the amateur musicians, there were nonetheless shared trends in the quality judgements of both groups. The findings from the experiment in Chapter 5 help clarify the effects of action-sound latency in DMIs and are of particular relevance to understanding how small discrepancies in timing behaviour can influence the perceived quality of an instrument and gestures of a performer. In relation to the projection model presented in Chapter 7, action-sound latency can be considered as effecting the reflection of action as it is projected through an instrument and the transparency of the lens responsible for this projection.

In terms of tangible experience action-sound latency was observed to impact on the 'feel' of the percussive DMI used in this experiment. On a number of occasions the effects of added latency were reported in relation to the feel of the instrument: increased weight being required to produce a sound; a variation to the *action* of the instrument from one that 'sounds immediately' (like a glockenspiel) to one that 'sounds late' (like a timpani); the sound shifting from being 'underneath the fingertips' and effortless, to disconnected and requiring concentration to play rhythmically. These reports highlight the *tangible effects* of latency in a DMI: as latency degrades the degree of control intimacy that a DMI contains, it impacts on the perceived feel and action of the instrument.

The findings from the experiment presented in Chapter 5 point to the prime importance of the stability of latency, alongside the importance of keeping it below a certain threshold (10ms and below). The recommendation here is that latency can be detrimental to expressive interaction with DMIs when a close coupling of action and sound is desired. The experiment presented in Chapter 5 suggests that having a fixed delay is better than one that fluctuates by even $\pm$3ms, and that the negative impact of this fluctuation holds across skill levels of musicians. In the case of the professional percussionists who took part in this experiment, a negative impact of 10ms latency was noted on their temporal accuracy, suggesting that even below this level there can be detrimental effects on performance when the rhythmic ability of the performer is particularly high.

### 8.2.2 *Haptic feedback*

The haptic feedback that a performer receives from an instrument is a further key component of the research presented in this thesis. As discussed in Section 2.1.4.2, when learning a musical instrument performers are highly reliant on the feedback that the instrument produces in response to their actions. It is through repeated practice of musical gestures that performers are able to build an internal model that couples action and perception, allowing them to make predictions about the outcomes of their movements and switch into a *feed-*

*forward* mode of interaction which moves much faster than a mode of performance reliant solely on feedback could achieve. In Section 2.3.1 I reviewed research that demonstrates the complexity of this feedback in musical performance, and the manner in which auditory and haptic feedback are differently relied upon depending on a performer's experience and on the instrument they are playing. In this thesis I have considered two types of haptic feedback present in an instrument's design: static (in the form of input modality), and dynamic (in the form of vibrotactile feedback).

#### 8.2.2.1 *Input modality*

The arrangement of an instrument's active parts and their physical characteristics contains a great deal of haptic information which serves to structure interaction and evoke patterns of action [80]. The static haptic feedback that an instrument produces is perceived by the performer through exploratory patterns of hand movement, as discussed in Section 2.1.1, and allows a performer to build an internal representation of an instrument. When coupled to sound production another layer is added to an instrument's behaviour: a touch sensor that produces no sound is tangibly different from one that is used to control sound, and the particular parameters under control can modulate the tangible characteristics of an instrument without changing its physical behaviour (see Section 2.3).

In the study presented in Chapter 6 we found that the physical supports that a particular input modality offer are important for encouraging patterns of pre-existing musical control, whether these supports are known in a sensorimotor sense or simply in a cultural sense. The physical presence of a set of strings was important for the guitarists that took part in this study, who displayed a preference for the technical familiarity of the strings, even when they were not behaving as they normally would on an acoustic instrument. For non-musicians there is less allegiance to a particular input modality, with their preferences shifting towards the cultural load of the guitar form regardless of input modality. By utilising familiar sensorimotor experiences through physical aspects of a DMI's design, mapping can become less of a digital question and more of a physical one.

In Section 3.4.2 I proposed that the concept of 'imageability' could apply to the formation of action-sound couplings that happen as a performer embodies a musical instrument. Static haptic feedback through physical design elements constitutes the basis of this idea, providing anchor points, landmarks, districts and paths that structure a performer's understanding of an instrument, but it is equally through the dynamic behaviour of these elements that tangible landmarks and guide points are formed.

#### 8.2.2.2 *Dynamic behaviour*

In the case of each of the instruments presented in Chapter 4 I observed subtly different temporal patterns of interaction under each feedback condition that relate to characteristics of the feedback. Each of the vibrotactile feedback conditions introduced a different set of push and pull effects to the instrument [355], changing the manner in which the performers touched the instrument in order to get most information from the feedback, and the manner in which they searched for certain notes. The vibrotactile feedback acts to reinforce structural elements in the instrument's design, as an additional piece of scaffolding for the performer to understand their interaction with an instrument.

From the reports of the participants that took part in the study presented in Chapter 4, it seems that reliability and determinacy of the behaviour of the feedback is an essential credential for it being integrated into performance. As the body is already doing many things when controlling an instrument, the findings from this study suggest that feedback should be relatively simple and task-focused, in order to allow for the prediction of an action's results. Perceptual attendance, as discussed in Section 2.3, also figures here: when the tuning-guidance of the tactile feedback was introduced to the instrument we noted that in certain cases tuning accuracy dropped due to a lack of concentration on the acoustic behaviour of the instrument. The vibrations became the centre of the performers' concentration, and they reported that significant effort was required in order to focus on both forms of feedback at once.

This type of feedback could be considered a kind of stabiliser to help in the training of sensorimotor control. As the performer gets a feel for the task, and for the distances between notes and the landmark points across the touch sensor, there is less need to rely on the supporting feedback. This could be seen as an additional case of 'complexity management' [284] as discussed in Section 3.5. In the first case of interaction with an unfamiliar DMI simplicity and predictability of feedback are of benefit, as there is only a certain bandwidth available for new learners. High variability between participants means that no one solution will suit everyone: some participants preferred complexity, some simplicity, and some didn't value the extra feedback at all. As with the sound that an instrument produces, there is a high degree of personal preference involved in the vibrational behaviour of a DMI.

### 8.2.3 *Summary*

This thesis has demonstrated that the tangibility of a DMI is a complex manifestation of many factors of an instrument's design. I have aimed to demonstrate how tangibility can relate to factors that are not usually discussed in relation to the sense of touch, including the immediacy with which an instrument responds to a performer's actions, the richness with which a performer's movements are reflected in the sound output of an instrument, and the characteristics of the haptic feedback that an instrument creates, whether static and manifested through physical aspects of its design, or dynamic in its vibrational behaviour in response to a performer's actions. Including these perspectives within an exposition of tangibility in DMI design will hopefully serve as inspiration for designers who aim to create rich and intimate musical instruments that target the full potential of performers' sensory and motor capabilities.

## 8.3 Statement of contribution

This research's main contribution is the demarcation of a field of enquiry – tangibility in the design of DMIs – that is currently underexplored in musical instrument design. Additionally, this research presents a series of design reflections that aim to inform current design practices and show the way for continued work in this area. Below is a reiteration of the main contributions of this thesis, in the order in which they have appeared:

- The vibrotactile feedback experiment in Chapter 4 demonstrates the degree to which musicians can exercise control of an instrument whilst attending to multimodal feedback.
- The effects of action-sound latency in DMIs found from the experiment in Chapter 5 are of particular relevance to understanding how small discrepancies in timing behaviour can influence the perceived quality of an instrument and the gestures of a performer.
- Findings relating to sensing richness and its role in judgements of instrument quality from the experiment presented in Chapter 6.
- Findings relating to static haptic feedback and physical form, and their influence on judgements of instrument quality from the experiment presented in Chapter 6.
- A series of reflections related to tangible aspects of DMI design outlined in Chapter 7.
- The projection model of performer-instrument interaction as outlined in Chapter 7, that can be of use in the evaluation, comparison and design of DMIs.
- A methodology that utilises technology probes in DMI research.
- A number of novel implementation strategies that involve the integration of dynamic tactile feedback into musical instruments,

design for low latency, and general guidelines for building self-contained and expressive DMIs.

## 8.4 methodological considerations

This thesis has employed a methodology that uses DMIs as technology probes. As noted by Hutchinson et al. [150], and as pointed out previously in this thesis, technology probes serve three functions: the social science goal of understanding the needs and desires of users, the engineering goal of field-testing the technology, and the design goal of inspiring users and researchers to think about new technologies. These goals closely reflect the kind of work this thesis sought to carry out in what concerns tangibility in performer-DMI interaction.

One of the benefits of the technology probe methodology is that it foregrounds the design of the device under study by recognising its provocative power in interaction. Rather than treating the device in an interaction study as neutral, the fine detail of how it was made and how it functions becomes part of the analysis through a critical reflection on design. Each of the probe instruments presented in this thesis were designed to enter into a musical scenario and provoke the performers to reflect on a certain area of DMI design. As technology probes, the instruments created during this research were not designed as finished or fully-featured instruments that would contribute to musical culture, but rather as opportunities to focus on a specific area of tangible interaction. Each of the instruments were made to gather information about performance in an invisible manner: the 'scientific instrumentation' of each of the studies was hidden within the DMI and functioned in the background.

As opposed to an 'in the wild' setting, the musical scenarios in which the instruments were tested were highly constructed, as were the musical tasks that the performers undertook. An effort was made in each study to leave space for free improvisation with the instrument before embarking on the musical tasks that were used as the basis for empirical analysis of the performers' performances. At the same time, the freedom that performers had with the instrument was always limited by the constraints of the study design, and an alternative approach could have been to involve fewer performers in a more longitudinal exploration of the instruments' potential. However, as the goal of much of the work presented in this thesis was to perform a type of musicological evaluation of the interaction between performer and instrument, a more widely comparative approach involving larger groups of performers, was more appropriate for the aims of this work.

Another important aspect of the methodology employed in this research is the role of surveys or questionnaires for gathering feedback from the performers in each study. These techniques are used extensively in user-testing in HCI, and necessarily take for granted a shared understanding between participants as to what particular terms or points of comparison mean. In the case of this thesis the terms were generally measures of quality, such as *responsiveness*, *temporal control*, *naturalness of interaction*. In order to mitigate differences in the interpretation of terms between participants I have included structured interviews throughout the studies that gather the opinions of the performers in their own words. This provides a much richer dataset from the participants, yet one that is more challenging to glean shared meaning and trends in opinion from. I have always aimed to temper the conclusions that I have drawn from the empirical ratings gathered throughout the studies with the first-hand reports from the participants that are scrutinised through thematic analysis. This adds complexity and nuance to the quantitative reports: musicians, just like any 'users' can differ greatly in their motivations, outlook and preferences.

Working with musicians has shaped the work presented in this thesis from various points of view. From an HCI perspective musicians can be considered as 'power users' of DMIs: the domain specificity of their extensive training and their high level of sensorimotor skill makes them the best-suited users for *extreme* testing of a DMI. This research has taken place in London, a city that hosts numerous orchestras, ensembles and conservatoires which have supplied me with highly skilled participants for my studies (this was particularly true for the orchestral percussionists involved in the study presented in Chapter 5). A side effect of working with such highly trained professionals is a relatively small sample size. In terms of statistical analysis, with such sample sizes it can be difficult to remove outliers or further group participants by shared trends in their ratings, without getting down to a participant by participant analysis. One place where statistical methods alone has been of great use is in the performance analysis conducted in Chapter 5. In this case I was able to gather data that could be successfully analysed using statistical techniques. In the studies presented in Chapter 4 and Chapter 6 drawing statistically meaningful conclusions was not as successful due to a number of factors related to the type of data gathered and sample sizes. This is discussed in each chapter.

Methodologically, new possibilities for quantitative analysis were granted by the technology used to construct the instruments. The Bela platform, the technological basis of all the instruments, was particularly important. It was due to the platform's high performance I was able to zoom in and investigate subtle areas of interaction with DMIs that are not easily accessible. As an example, the study presented in Chapter 5 had the possibility of covering new ground thanks to the high-resolution and low-latency performance of the platform. Working with this technology also allowed me to create instruments that are at once highly responsive and sonically interesting, while being able to behave like a piece of high-end lab equipment in their ability to record performer interaction. The actual design work of each of the instruments in this thesis has had an important influence on my findings in each case and I hope that through the detailed information I have given about the build of each instrument I have been able to represent them clearly in their functioning and purpose (see Section 4.3, Section 5.3 and Section 6.3).

This thesis has covered a broad range of topics related to tangible experience with DMIs, and each of the topics could be the subject of further investigation, and my hope is that this wide-ranging perspective can open up opportunities for future work.

## 8.5 FUTURE DIRECTIONS

The main goal of this thesis has been to contribute to a discourse on touch in DMI design and to nuance our understanding of tangibility and physicality with regards to such instruments. Accordingly this research seeks to encourage other designers, particularly those who are working with sound, musical or not, to reflect more on issues of tangibility and richness in the design process. The main themes that this research investigates (control intimacy, latency and feedback) are fundamental to the design of DMIs and do not necessarily need to be framed as connected to haptic experience; I have chosen to do so however, in order to refine what we mean when we talk about touch, providing some case studies that demonstrate these factors in practice.

While each of the studies presented in this research represent a defined and at times very field-specific contribution (to music psychology [158], tangible computing [155], and DMI design [157, 159]), they also each open up territory for future research. Using similar experimental techniques as those introduced in Chapter 5 it would be possible to conduct a series of studies on the impact of latency on different instrument types and with different groups of performers. An example of such a study, which would combine the design work from Chapter 5 and Chapter 6, would be to address the impact of latency on guitar techniques such as strumming and fingerpicking. Insightful findings could also be granted by an exploration of control intimacy and latency together, an area of investigation that fell outside of the limits of this thesis: for example a study could address the different effects of latency when an instrument is using audio-driven synthesis versus sample-triggering. Further applications of the techniques used in Chapter 4 could also be explored in an interactive device that has less time constraints than a musical instrument.

Each of the studies presented in this thesis has observed the 'first contact' that a performer makes with a DMI. In each case the encounter lasted just over an hour and hence all of the findings presented in this thesis relate to performers' first impressions of the instruments, and to an initial testing out of the instruments' potential. Given a more long-term engagement with each of the instruments different patterns of interaction would inevitably start to appear. As an anecdotal example I have personally known cases where musicians are able to adapt to levels of latency in an instrument that others would find crippling. Longitudinal evaluations of each of these tangible aspects could produce interesting results, for example an analysis of how long it takes performers to incorporate different types of feedback into their existing practice, or a study of the influence of gradually increasing control intimacy.

The research presented in this thesis has made use of a hybrid methodology that draws on techniques from HCI, from design and from musicology. Due to the complexity of the musical context and the technological, cultural, and aesthetic threads that it brings together, the probe approach can act as a suitable alternative to more common HCI techniques of user-testing, and it could potentially be of use to others working on research that relates to DMIs and on interaction with digital systems in general. The technology probe approach is advantageous because it does not underestimate the power of the design of a device, rather it makes space for an acknowledgement of the role of design, of the designer, and of the manner in which designs can provoke and influence people who interact with them. Additionally, the model presented in Chapter 7 can be applied to the evaluation and comparison of DMI designs and I hope that this will also encourage further discussion around the embodied nature of interaction with a musical instrument.

There is still great potential for collaborations between the fields of tangible and physical computing and that of new musical instrument design. Many researchers have been involved in both fields since their beginnings and there are many areas of crossover and mutual agreement, however both fields have valuable lessons to share with each other. From tangible computing there are lessons about the emotive power of touch, the realm of control metaphors, and the meaning contained in materials and physical action. From digital musical instrument design there are lessons about the complexity of sensorimotor skill development, intimate and nuanced control, and tangible interaction in service of producing a product that is in most part intangible. This thesis has aimed to bring discussions from both fields a little closer together, as will my future work.

## 8.6 CONCLUDING REMARKS

Musical instruments are amongst the most complex human artefacts. Beyond their primary function as translators of action into sound, musical instruments provide the basis of musical thinking, and are the result of countless physical and cultural considerations which in turn open up a multitude of compelling areas of research. By studying and designing musical instruments we can understand a great deal about our cognitive and motor faculties, about how we derive meaning from sound, and about how skill forms and interacts with technological artefacts.

Digital technologies represent but another evolution in the long history of musical instrument design, one which opens up unprecedented possibilities but also complicates our relationship to musical instruments and to musical thinking in unprecedented ways. This thesis has engaged with this double-edged sword of DMI design, investigating the barriers that DMIs pose to the rich physical experience of performing with an instrument and attempting to break down some of those barriers or to imagine them differently. As digital musical instruments increase in their sophistication, an increased appreciation of the tangible aspects of interaction is an inevitability. In the meantime, with the technological limitations of the moment, the work presented here has weighed up the sensorimotor complexity of musical interaction against the constraints of design and put forward proposals that will hopefully be useful to thinking about, building and playing rich DMIs in the future.

# BIBLIOGRAPHY


[1] R Abrams. "Adventures in tangible computing: The work of interaction designer 'Durrell Bishop'in context." In: *magistrska naloga, Royal College of Art, London* (1999).

[2] Bernard D Adelstein, Durand R Begault, Mark R Anderson, and Elizabeth M Wenzel. "Sensitivity to haptic-audio asynchrony." In: *Proceedings of the 5th international conference on Multimodal interfaces*. ACM. 2003, pp. 73–76.

[3] Marko Aho. *The Tangible in Music: The Tactile Learning of a Musical Instrument*. Routledge, 2016.

[4] Robert S Allison, Laurence R Harris, Michael Jenkin, Urszula Jasiobedzka, and James E Zacher. "Tolerance of temporal delay in virtual environments." In: *Virtual Reality, 2001. Proceedings. IEEE*. IEEE. 2001, pp. 247–254.

[5] Kristina Andersen and Nicholas Ward. "Learning from the crackle exhibition." In: *Proceedings of the Tenth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM. 2017, pp. 225–231.

[6] John R Anderson. "Acquisition of cognitive skill." In: *Psychological review* 89.4 (1982), p. 369.

[7] Michelle Annett, Albert Ng, Paul Dietz, Walter F Bischof, and Anoop Gupta. "How low should we go?: understanding the perception of latency while inking." In: *Proceedings of Graphics Interface 2014*. Canadian Information Processing Society. 2014, pp. 167–174.

[8] Newton Armstrong. "An enactive approach to digital musical instrument design." PhD thesis. Princeton University, 2006.

[9] Gisa Aschersleben. "Temporal control of movements in sensorimotor synchronization." In: *Brain and cognition* 48.1 (2002), pp. 66–79.

[10] Gisa Aschersleben and Wolfgang Prinz. "Delayed auditory feedback in synchronization." In: *Journal of motor behavior* 29.1 (1997), pp. 35–46.

[11] Anders Askenfelt and Erik V Jansson. "From touch to string vibrations - the initial course of the piano tone." In: *Dept. for Speech Music and Hearing, Quarterly Progress and Status Report* 29.1 (1988), pp. 31–109.

[12] Anders Askenfelt and Erik V Jansson. "On vibration sensation and finger touch in stringed instrument playing." In: *Music Perception: An Interdisciplinary Journal* 9.3 (1992), pp. 311–349.

[13] Rimas Avizienis, Adrian Freed, Takahiko Suzuki, and David Wessel. "Scalable Connectivity Processor for Computer Music Performance Systems." In: *Proc. International Computer Music Conference (ICMC)*. 2000.

[14] Anant Baijal, Julia Kim, Carmen Branje, Frank Russo, and Deborah I Fels. "Composing vibrotactile music: A multi-sensory experience with the emoti-chair." In: *Haptics Symposium (HAPTICS), 2012 IEEE*. IEEE. 2012, pp. 509–515.

[15] John Baily et al. "Ethnomusicology, intermusability, and performance practice." In: *The new (ethno) musicologies* 8 (2008), p. 117.

[16] Roland Barthes. *Image-music-text*. Macmillan, 1978.

[17] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. "Fitting Linear Mixed-Effects Models Using lme4." In: *Journal of Statistical Software* 67.1 (2015), pp. 1–48.

[18] Eliot Bates. "The Social Life of Musical Instruments." In: *Ethnomusicology* 56.1895 (2012), pp. 363–395.

[19] Francis Bebey. *African music: A people's art*. Chicago Review Press, 1999.

[20] Klaus-Ernst Behne and Clemens Wöllner. "Seeing or hearing the pianists? A synopsis of an early audiovisual perception experiment and a replication." In: *Musicae Scientiae* 15.3 (2011), pp. 324–342.

[21] Steve Benford, Holger Schnädelbach, Boriana Koleva, Rob Anastasi, Chris Greenhalgh, Tom Rodden, Jonathan Green, Ahmed Ghali, Tony Pridmore, Bill Gaver, et al. "Expected, sensed, and desired: A framework for designing sensing-based interaction." In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 12.1 (2005), pp. 3–30.

[22] Steve Benford, Adrian Hazzard, Alan Chamberlain, Kevin Glover, Chris Greenhalgh, Liming Xu, Michaela Hoare, and Dimitrios Darzentas. "Accountable artefacts: the case of the Carolan guitar." In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM. 2016, pp. 1163–1175.

[23] Peter Bennett and Sile O'Modhrain. "The BeatBearing: a tangible rhythm sequencer." In: *Proc. of NordiCHI*. 2008.

[24] Peter Bennett, Nicholas Ward, Sile O'Modhrain, and Pedro Rebelo. "DAMPER: a platform for effortful interface development." In: *Proceedings of the 7th International Conference on New Interfaces for Musical Expression*. ACM. 2007, pp. 273–276.

[25] Edgar Berdahl and Alexandros Kontogeorgakopoulos. "The firefader design: Simple, open-source, and reconfigurable haptics for musicians." In: *Proceedings of the 9th Sound and Music Computing Conference*. 2012, pp. 90–98.

[26] Edgar Berdahl, Günter Niemeyer, and Julius O Smith. "Using Haptics to Assist Performers in Making Gestures to a Musical Instrument." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2009, pp. 177–182.

[27] Edgar Berdahl, Hans-Christoph Steiner, and Collin Oldham. "Practical Hardware and Algorithms for Creating Haptic Musical Instruments." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2008, pp. 61–66.

[28] Soumitra Bhat. "TouchTone: an electronic musical instrument for children with hemiplegic cerebral palsy." In: *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*. ACM. 2010, pp. 305–306.

[29] Karin Bijsterveld and Marten Schulp. "Breaking into a world of perfection: Innovation in today's classical musical instruments." In: *Social Studies of Science* 34.5 (2004), pp. 649–674.

[30] David Birnbaum and Marcelo M Wanderley. "A Systematic Approach to Musical vibrotactile feedback." In: *Proc. International Computer Music Conference (ICMC)*. 2007.

[31] David Birnbaum, Rebecca Fiebrink, Joseph Malloch, and Marcelo M Wanderley. "Towards a dimension space for musical devices." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. National University of Singapore. 2005, pp. 192–195.

[32] John W Black. "The effect of delayed side-tone upon vocal rate and intensity." In: *Journal of Speech and Hearing Disorders* 16.1 (1951), pp. 56–60.

[33] Kirsten Boehner, Janet Vertesi, Phoebe Sengers, and Paul Dourish. "How HCI interprets the probes." In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2007, pp. 1077–1086.

[34] Kenneth R Boff, Lloyd Kaufman, and James P Thomas. *Handbook of perception and human performance*. Vol. 2. Wiley New York, 1986.

[35] Jay David Bolter and Diane Gromala. *Windows and mirrors: Interaction design, digital art, and the myth of transparency*. MIT press, 2003.

[36] AJ Bongers. "Tactual display of sound properties in electronic musical instruments." In: *Displays* 18.3 (1998), pp. 129–133.

[37] Bert Bongers. "The use of active tactile and force feedback in timbre controlling electronic instruments." In: *Proceedings of the International Computer Music Conference*. 1994, pp. 171–171.

[38] John Bowers and Sten Olof Hellström. "Simple interfaces to complex sound in improvised music." In: *CHI'00 extended abstracts on Human factors in computing systems*. ACM. 2000, pp. 125–126.

[39] Geoffrey C Bowker. *Memory practices in the sciences*. Vol. 205. MIT Press, Cambridge, MA, 2005.

[40] Virginia Braun and Victoria Clarke. "Using thematic analysis in psychology." In: *Qualitative research in psychology* 3.2 (2006), pp. 77–101.

[41] William Brent. "Aspects of gesture in digital musical instrument design." In: *Proc. International Computer Music Conference (ICMC)*. 2011.

[42] Jean-Pierre Bresciani, Marc O Ernst, Knut Drewing, Guillaume Bouyer, Vincent Maury, and Abderrahmane Kheddar. "Feeling what you hear: auditory signals can modulate tactile tap perception." In: *Experimental brain research* 162.2 (2005), pp. 172–180.

[43] PL Brooks, BJ Frost, JL Mason, and DM Gibson. "Continuing evaluation of the Queen's University tactile vocoder. I: Identification of open set words." In: *Journal of rehabilitation research and development* 23.1 (1986), pp. 119–128.

[44] Dom Brown, Chris Nash, and Tom Mitchell. "A user experience review of music interaction evaluations." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2017.

[45] Judith C Brown and Kathryn V Vaughn. "Pitch center of stringed instrument vibrato tones." In: *The Journal of the Acoustical Society of America* 100.3 (1996), pp. 1728–1735.

[46] W Buxton. "There's more to interaction than meets the eye: Some issues in manual input." In: *Human-computer interaction*. Morgan Kaufmann Publishers Inc. 1987, pp. 366–375.

[47] Claude Cadoz. "Supra-instrumental interactions and gestures." In: *Journal of New Music Research* 38.3 (2009), pp. 215–230.

[48] Claude Cadoz and Marcelo M Wanderley. *Gesture-music*. 2000.

[49] Claude Cadoz, Annie Luciani, Jérôme Villeneuve, Alexandros Kontogeorgakopoulos, and Iannis Zannos. "Tangibility, presence, materiality, reality in artistic creation with digital technology." In: *40th International Computer Music Conference/11th Sound and Music Computing Conference*. 2014, pp. 754–761.

[50] Gemma Calvert, Charles Spence, and Barry E Stein. *The handbook of multisensory processes*. MIT press, 2004.

[51] Daniel J Cameron and Jessica A Grahn. "Enhanced timing abilities in percussionists generalize to rhythms without a musical beat." In: *Frontiers in human neuroscience* 8 (2014), p. 1003.

[52] Baptiste Caramiaux, Patrick Susini, Tommaso Bianco, Frédéric Bevilacqua, Olivier Houix, Norbert Schnell, and Nicolas Misdariis. "Gestural Embodiment of Environmental Sounds: an Experimental Study." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. Vol. 11. 2011.

[53] Tom Carter, Sue Ann Seah, Benjamin Long, Bruce Drinkwater, and Sriram Subramanian. "UltraHaptics: multi-point mid-air haptic feedback for touch surfaces." In: *Proceedings of the 26th annual ACM symposium on User interface software and technology*. ACM. 2013, pp. 505–514.

[54] Nicolas Castagné, Claude Cadoz, Jean-Loup Florens, and Annie Luciani. "Haptics in computer music: a paradigm shift." In: *EuroHaptics, Munich, Germany*. 2004, pp. 5–7.

[55] Chris Chafe. "Tactile audio feedback." In: *Proceedings of the International Computer Music Conference*. 1993, pp. 76–76.

[56] Chris Chafe and Michael Gurevich. "Network time delay and ensemble accuracy: Effects of latency, asymmetry." In: *Audio Engineering Society Convention 117*. 2004.

[57] Jessie Chen, Marjorie H Woollacott, Steven Pologe, and George P Moore. "Pitch and space maps of skilled cellists: accuracy, variability, and error correction." In: *Experimental brain research* 188.4 (2008), pp. 493–503.

[58] Insook Choi. "Cognitive engineering of gestural primitives for multi-modal interaction in a virtual environment." In: *Systems, Man, and Cybernetics, 1998. 1998 IEEE International Conference on*. Vol. 2. IEEE. 1998, pp. 1101–1106.

[59] Eric Clarke. *Ways of listening: An ecological approach to the perception of musical meaning*. Oxford University Press, 2005.

[60] Carter Compton Collins. "Tactile television-mechanical and electrical image projection." In: *IEEE Transactions on man-machine systems* 11.1 (1970), pp. 65–71.

[61] Perry Cook. "Principles for designing computer music controllers." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2001.

[62] Arnie Cox. "The mimetic hypothesis and embodied musical meaning." In: *Musicae scientiae* 5.2 (2001), pp. 195–212.

[63] Richard Coyne, Martin Parker, and Pedro Rebelo. "Resisting the seamless interface." In: *International Journal of Architectural Computing* 2.4 (2004), pp. 429–442.

[64] Mihaly Csikszentmihalyi. *Flow: The psychology of optimal performance*. 1990.

[65] Sofia Dahl. "The playing of an accent–preliminary observations from temporal and kinematic analysis of percussionists." In: *Journal of New Music Research* 29.3 (2000), pp. 225–233.

[66] Sofia Dahl and Roberto Bresin. "Is the player more influenced by the auditory than the tactile feedback from the instrument?" In: *Proc. of the COST-G6 Workshop on Digital Audio Effects (DAFx-01), Limerick*. 2001, pp. 194–197.

[67] Helios De Rosario-Martinez. *phia: Post-Hoc Interaction Analysis*. R package version 0.2-1. 2015. URL: https://CRAN.R-project.org/package=phia.

[68] Jonathan De Souza. *Music at hand: Instruments, bodies, and cognition*. Oxford University Press, 2017.

[69] Jonathan Deber, Ricardo Jota, Clifton Forlines, and Daniel Wigdor. "How Much Faster is Fast Enough?" In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15* 1 (2015), pp. 1827–1836.

[70] François Delalande. "Le Geste, outil d'analyse: quelques enseignements d'une recherche sur la gestique de Glenn Gould." In: *Analyse musicale* 10 (1988), pp. 43–46.

[71] Stefano Delle Monache, Stefano Papetti, Pietro Polotti, and Davide Rocchesso. "Sonically augmented found objects." In: (2008), 154—157.

[72] Tom Djajadiningrat, Stephan Wensveen, Joep Frens, and Kees Overbeeke. "Tangible products: redressing the balance between appearance and action." In: *Personal and Ubiquitous Computing* 8.5 (2004), pp. 294–309.

[73] Christopher Dobrian and Daniel Koppelman. "The'E'in NIME: musical expression with new computer interfaces." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. IRCAM—Centre Pompidou. 2006, pp. 277–282.

[74] Mine Doğantan-Dack. "In the beginning was gesture: Piano touch and the phenomenology of the performing body." In: *New perspectives on music and gesture* (2011), pp. 243–265.

[75] Paul Dourish. "Embodied interaction: Exploring the foundations of a new approach to HCI." In: *Unpublished paper, online: http://www. ics. uci. edu/˜ jpd/publications/misc/embodied. pdf* (1999).

[76] Paul Dourish. *Where the action is: the foundations of embodied interaction*. MIT press, 2004.

[77] Nathalie Ehrlé and Séverine Samson. "Auditory discrimination of anisochrony: Influence of the tempo and musical backgrounds of listeners." In: *Brain and Cognition* 58.1 (2005), pp. 133–147.

[78] Zohar Eitan and Inbar Rothschild. "How music touches: Musical parameters and listeners' audio-tactile metaphorical mappings." In: *Psychology of Music* 39.4 (2011), pp. 449–467.

[79] Marc O Ernst and Heinrich H Bülthoff. "Merging the senses into a robust percept." In: *Trends in cognitive sciences* 8.4 (2004), pp. 162–169.

[80] Georg Essl and Sile O'Modhrain. "An enactive approach to the design of new tangible musical instruments." In: *Organised Sound* 11.03 (2006), pp. 285–296.

[81] F Alton Everest. *Master handbook of acoustics*. 2001.

[82] Udo Eversheim and Otmar Bock. "Evidence for processing stages in skill acquisition: a dual-task study." In: *Learning & Memory* 8.4 (2001), pp. 183–189.

[83] Sidney Fels. "Designing for intimacy: Creating new interfaces for musical expression." In: *In Proc. of the IEEE* 92.4 (2004), pp. 672–685.

[84] Sidney Fels, Ashley Gadd, and Axel Mulder. "Mapping transparency through metaphor: towards more expressive musical instruments." In: *Organised Sound* 7.2 (2002), pp. 109–126.

[85] Ylva Fernaeus and Petra Sundström. "The material move how materials matter in interaction design research." In: *Proceedings of the Designing Interactive Systems Conference*. ACM. 2012, pp. 486–495.

[86] Steven A Finney. "Auditory feedback and musical keyboard performance." In: *Music Perception: An Interdisciplinary Journal* 15.2 (1997), pp. 153–174.

[87] Paul Morris Fitts and Michael I Posner. *Human performance*. Belmont, CA: Brooks/Cole, 1967.

[88] George W Fitzmaurice, Hiroshi Ishii, and William AS Buxton. "Bricks: laying the foundations for graspable user interfaces." In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co. 1995, pp. 442–449.

[89] Jean-Loup Florens, Annie Luciani, Claude Cadoz, and Nicolas Castagné. "ERGOS: Multi-degrees of freedom and versatile force-feedback panoply." In: *EuroHaptics 2004*. 2004.

[90] Federico Fontana, Federico Avanzini, Hanna Järveläinen, Stefano Papetti, Francesco Zanini, and Valerio Zanini. "Perception of interactive vibrotactile cues on the acoustic grand and upright piano." In: *Proc. International Computer Music Conference (ICMC)*. 2014.

[91] Federico Fontana, Hanna Järveläinen, Stefano Papetti, Federico Avanzini, Giorgio Klauer, Lorenzo Malavolta, Conservatorio di Musica, and Cesare Pollini. "Rendering and subjective evaluation of real vs. synthetic vibrotactile cues on a digital piano keyboard." In: *Proc. of the Sound and Music Computing Conference 2015, Maynooth, Ireland*. 2015.

[92] Karmen Franinović and Stefania Serafin. *Sonic interaction design*. MIT Press, 2013.

[93] Claudia Fritz, Jacques Poitevineau, and Stephen McAdams. "Investigating the role of auditory and tactile modalities in violin quality evaluation." In: *PloS one* 9.12 (2014), e112552.

[94] Shinya Fujii, Masaya Hirashima, Kazutoshi Kudo, Tatsuyuki Ohtsuki, Yoshihiko Nakamura, and Shingo Oda. "Synchronization Error of Drum Kit Playing with a Metronome at Different Tempi by Professional Drummers." In: *Music Perception: An Interdisciplinary Journal* 28.5 (2011), pp. 491–503.

[95] Waka Fujisaki and Shin'ya Nishida. "Audio–tactile superiority over visuo–tactile and audio–visual combinations in the temporal resolution of synchrony perception." In: *Experimental brain research* 198.2-3 (2009), pp. 245–259.

[96] Robert Fulford, Jane Ginsborg, and Juliet Goldbart. "Learning not to listen: the experiences of musicians with hearing impairments." In: *Music Education Research* 13.4 (2011), pp. 447–464.

[97] A Cavan Fyans, Michael Gurevich, and Paul Stapleton. "Examining the Spectator Experience." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. Vol. 10. 2010, pp. 1–4.

[98] Bruno Galantucci, Carol A Fowler, and Michael T Turvey. "The motor theory of speech perception reviewed." In: *Psychonomic bulletin & review* 13.3 (2006), pp. 361–377.

[99] Alexander Galembo and Anders Askenfelt. "Quality assessment of musical instruments-effects of multimodality." In: *5th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM5), Hannover, Germany*. 2003.

[100] Alberto Gallace and Charles Spence. *In touch with the future: The sense of touch from cognitive neuroscience to virtual reality*. Oxford University Press, Oxford, 2014.

[101] Alberto Gallace, Hong Z Tan, and Charles Spence. "Can tactile stimuli be subitised? An unresolved controversy within the literature on numerosity judgments." In: *Perception* 37.5 (2008), pp. 782–800.

[102] Jérémie Garcia, Theophanis Tsandilas, Carlos Agon, and Wendy E Mackay. "Structured observation with polyphony: a multifaceted tool for studying music composition." In: *Proceedings of the 2014 conference on Designing interactive systems*. ACM. 2014, pp. 199–208.

[103] Anne Gates, John L Bradshaw, and Norman C Nettleton. "Effect of different delayed auditory feedback intervals on a music performance task." In: *Perception & Psychophysics* 15.1 (1974), pp. 21–25.

[104] Bill Gaver, Tony Dunne, and Elena Pacenti. "Design: cultural probes." In: *interactions* 6.1 (1999), pp. 21–29.

[105] William W Gaver. "Technology affordances." In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 1991, pp. 79–84.

[106] William W Gaver. "What in the world do we hear?: An ecological approach to auditory event perception." In: *Ecological psychology* 5.1 (1993), pp. 1–29.

[107] AM Gentile. "Skill acquisition: Action, movement, and neuromotor processes." In: *Movement science* (2000), pp. 111–187.

[108] C Gerard and M Rosenfeld. "Musical expertise and temporal regulation." In: *Annee Psychologique* 95.4 (1995), pp. 571–591.

[109] John M Geringer, Rebecca B MacLeod, and Julia C Ellis. "Two studies of pitch in string instrument vibrato: Perception and pitch matching responses of university and high school string players." In: *International journal of music education* 32.1 (2014), pp. 19–30.

[110] George A Gescheider. "Resolving of successive clicks by the ears and skin." In: *Journal of experimental psychology* 71.3 (1966), p. 378.

[111] James J Gibson. "Observations on active touch." In: *Psychological review* 69.6 (1962), p. 477.

[112] James J Gibson. *The senses considered as perceptual systems.* Boston: Houghton Mifflin, 1966.

[113] James J Gibson. *The Ecological Approach to Visual Perception*. Psychology Press, 1986.

[114] Marcello Giordano and Marcelo M Wanderley. "Perceptual and technological issues in the design of vibrotactile-augmented interfaces for music technology and media." In: *International Workshop on Haptic and Audio Interaction Design*. Springer. 2013, pp. 89–98.

[115] Marcello Giordano and Marcelo M Wanderley. "Follow the tactile metronome: vibrotactile stimulation for tempo synchronization in music performance." In: *Proc. of the Sound and Music Computing Conference, Maynooth, Ireland*. 2015.

[116] Rolf Inge Godøy. "Motor-mimetic music cognition." In: *Leonardo* 36.4 (2003), pp. 317–319.

[117] Rolf Inge Godøy. "Understanding coarticulation in musical experience." In: *International Symposium on Computer Music Modeling and Retrieval*. Springer. 2013, pp. 535–547.

[118] Rolf Inge Godøy, Egil Haga, and Alexander Refsum Jensenius. "Playing "air instruments": mimicry of sound-producing gestures by novices and experts." In: *International Gesture Workshop*. Springer. 2005, pp. 256–267.

[119] Rolf Inge Godøy and Marc Leman. *Musical gestures: Sound, movement, and meaning*. Routledge, 2010.

[120] Rolf Inge Godøy, Minho Song, Kristian Nymoen, Mari Romarheim Haugen, and Alexander Refsum Jensenius. "Exploring sound-motion similarity in musical experience." In: *Journal of New Music Research* 45.3 (2016), pp. 210–222.

[121] Werner Goebl and Caroline Palmer. "Tactile feedback and timing accuracy in piano performance." In: *Experimental Brain Research* 186.3 (2008), pp. 471–479.

[122] Werner Goebl and Caroline Palmer. "Tactile feedback and timing accuracy in piano performance." In: *Experimental Brain Research* 186.3 (2008), pp. 471–479.

[123] Steve Guest, Caroline Catmur, Donna Lloyd, and Charles Spence. "Audiotactile interactions in roughness perception." In: *Experimental Brain Research* 146.2 (2002), pp. 161–171.

[124] Eric Gunther and Sile O'Modhrain. "Cutaneous grooves: composing for the sense of touch." In: *Journal of New Music Research* 32.4 (2003), pp. 369–381.

[125] Michael Gurevich, Paul Stapleton, and Adnan Marquez-Borbon. "Style and Constraint in Electronic Musical Instruments." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2010.

[126] Michael Gurevich and Jeffrey Treviño. "Expression and its discontents: toward an ecology of musical creation." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. ACM. 2007, pp. 106–111.

[127] Patrick Haggard and J Randall Flanagan. *Hand and brain: the neurophysiology and psychology of hand movements*. Academic Press, 1996.

[128] Sarah-Indriyati Hardjowirogo. "Instrumentality. On the construction of instrumental identity." In: *Musical instruments in the 21st century*. Springer, 2017, pp. 9–24.

[129] Vanessa Harrar and Charles Spence. "The taste of cutlery: how the taste of food is affected by the weight, size, shape, and colour of the cutlery used to eat it." In: *Flavour* 2.1 (2013), p. 21.

[130] Jacob Harrison, Robert H Jack, Fabio Morreale, and Andrew McPherson. "When is a Guitar not a Guitar? Cultural Form, Input Modality and Expertise." In: *Proceedings of the international conference on new interfaces for musical expression*. 2018.

[131] Steve Harrison, Deborah Tatar, and Phoebe Sengers. "The three paradigms of HCI." In: *Alt. Chi. Session at the SIGCHI Conference on Human Factors in Computing Systems San Jose, California, USA*. 2007, pp. 1–18.

[132] Phillip M Hash. "Teaching Instrumental Music to Deaf and Hard of Hearing Students." In: *Research and Issues in Music Education* 1.1 (2003), n1.

[133] Lauren Hayes. "Vibrotactile Feedback-Assisted Performance." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2011, pp. 72–75.

[134] Vincent Hayward. "A brief taxonomy of tactile illusions and demonstrations that can be done in a hardware store." In: *Brain research bulletin* 75.6 (2008), pp. 742–752.

[135] Vincent Hayward, Oliver R Astley, Manuel Cruz-Hernandez, Danny Grant, and Gabriel Robles-De-La-Torre. "Haptic interfaces and devices." In: *Sensor Review* 24.1 (2004), pp. 16–29.

[136] Martin Heidegger. *The question concerning technology*. Harper & Row New York, 1977.

[137] Kahl Hellmer and Guy Madison. "Quantifying microtiming patterning and variability in drum kit recordings." In: *Music Perception: An Interdisciplinary Journal* 33.2 (2015), pp. 147–162.

[138] John R Hennessy. "Cutaneous sensitivity communications." In: *Human factors* 8.5 (1966), pp. 463–469.

[139] Cristy Ho and Charles Spence. "Affective multisensory driver interface design." In: *International Journal of Vehicle Noise and Vibration* 9.1-2 (2013), pp. 61–74.

[140] Oliver Hödl and Geraldine Fitzpatrick. "Exploring the design space of hand-controlled guitar effects for live music." In: *Proc. International Computer Music Conference (ICMC)*. 2013.

[141] Bernhard Hommel. "Acquisition and control of voluntary action." In: *Voluntary action: Brains, minds, and sociality* (2003), pp. 34–48.

[142] Mantle Hood. *The ethnomusicologist*. Vol. 140. Kent State Univ Pr, 1982.

[143] Michael S Horn. "The role of cultural forms in tangible interaction design." In: *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*. ACM. 2013, pp. 117–124.

[144] Eva Hornecker. "Beyond affordance: tangibles' hybrid nature." In: *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*. ACM. 2012, pp. 175–182.

[145] Eva Hornecker and Jacob Buur. "Getting a grip on tangible interaction: a framework on physical space and social interaction." In: *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM. 2006, pp. 437–446.

[146] Kirsten Hötting and Brigitte Röder. "Hearing cheats touch, but less in congenitally blind than in sighted individuals." In: *Psychological Science* 15.1 (2004), pp. 60–64.

[147] Peter Howell. "A model of timing interference to speech control in normal and altered listening conditions applied to the treatment of stuttering." In: *Speech motor control in normal and disordered speech* (2001), pp. 291–294.

[148] Andy Hunt, Marcelo M Wanderley, and Matthew Paradis. "The importance of parameter mapping in electronic instrument design." In: *Journal of New Music Research* 32.4 (2003), pp. 429–440.

[149] David Huron and Jonathon Berec. "Characterizing idiomatic organization in music: A theory and case study of musical affordances." In: (2009).

[150] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, et al. "Technology probes: inspiring design for and with families." In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2003, pp. 17–24.

[151] Hiroshi Ishii and Brygg Ullmer. "Tangible bits: towards seamless interfaces between people, bits and atoms." In: *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*. ACM. 1997, pp. 234–241.

[152] Hiroshi Ishii, Carlo Ratti, Ben Piper, Yao Wang, Assaf Biderman, and Eran Ben-Joseph. "Bringing clay and sand into digital design—continuous tangible user interfaces." In: *BT technology journal* 22.4 (2004), pp. 287–299.

[153] Robert H Jack, Andrew McPherson, and Tony Stockman. "Designing Tactile Musical Devices with and for Deaf Users: a Case Study." In: *Proceedings of the International Conference on the Multimodal Experience of Music*. HRI Online Publications. 2015.

[154] Robert H. Jack, Tony Stockman, and Andrew McPherson. "Effect of latency on performer interaction and subjective quality assessment of a digital musical instrument." In: *Proceedings of the Audio Mostly 2016*. ACM. 2016, pp. 116–123.

[155] Robert H Jack, Tony Stockman, and Andrew McPherson. "Navigation of Pitch Space on a Digital Musical Instrument with Dynamic Tactile Feedback." In: *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM. 2016, pp. 3–11.

[156] Robert H Jack, Tony Stockman, and Andrew McPherson. "Maintaining and Constraining Performer Touch in the Design of Digital Musical Instruments." In: *Proceedings of the TEI'17: Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*. ACM. 2017, pp. 717–720.

[157] Robert H Jack, Tony Stockman, and Andrew McPherson. "Rich gesture, reduced control: the influence of constrained mappings on performance technique." In: *4th International Conference on Movement Computing (MOCO'17), London, United Kingdom*. 2017.

[158] Robert H Jack, Adib Mehrabi, Tony Stockman, and Andrew McPherson. "Action-sound Latency and the Perceived Quality of Digital Musical Instruments: Comparing Professional Percussionists and Amateur Musicians." In: *Music Perception: An Interdisciplinary Journal* (2018).

[159] Robert H Jack, Jacob Harrison, Fabio Morreale, and Andrew McPherson. "Democratising DMIs: the Relationship of Expertise and Control Intimacy." In: *Proceedings of the international conference on new interfaces for musical expression*. 2018.

[160] David A Jaffe and Julius O Smith. "Extensions of the Karplus-Strong plucked-string algorithm." In: *Computer Music Journal* 7 (1983), pp. 56–69.

[161] Arvid Jense and Berry Eggen. "Awakening the Synthesizer Knob: Gestural Perspectives." In: *Machines* 3.4 (2015).

[162] Alexander Refsum Jensenius. "To gesture or not? an analysis of terminology in NIME proceedings 2001-2013." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2014.

[163] Alexander Refsum Jensenius, Marcelo M Wanderley, Rolf Inge Godøy, and Marc Leman. "Musical Gestures: concepts and methods in research." In: *Musical gestures: Sound, movement, and meaning*. Ed. by Rolf Inge Godøy and Marc Leman. Routledge, 2009, pp. 12–35.

[164] Roland S Johansson and J Randall Flanagan. "Coding and use of tactile signals from the fingertips in object manipulation tasks." In: *Nature reviews. Neuroscience* 10.5 (2009), p. 345.

[165] Lynette A Jones and Nadine B Sarter. "Tactile displays: Guidance for their design and application." In: *Human factors* 50.1 (2008), pp. 90–111.

[166] Randy Jones, Peter F Driessen, W Andrew Schloss, and George Tzanetakis. "A Force-Sensitive Surface for Intimate Control." In: *NIME*. Citeseer. 2009, pp. 236–241.

[167] Sergi Jordà. "Instruments and players: Some thoughts on digital lutherie." In: *Journal of New Music Research* 33.3 (2004), pp. 321–341.

[168] Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. "The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces." In: *Proceedings of the 1st international conference on Tangible and embedded interaction*. ACM. 2007, pp. 139–146.

[169] Ricardo Jota, Albert Ng, Paul Dietz, and Daniel Wigdor. "How fast is fast enough?: a study of the effects of latency in direct-touch pointing tasks." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2013, pp. 2291–2300.

[170] Marja-Leena Juntunen and Leena Hyvönen. "Embodiment in musical knowing: how body movement facilitates learning within Dalcroze Eurhythmics." In: *British Journal of Music Education* 21.2 (2004), pp. 199–214.

[171] T. Kaaresoja, E. Anttila, and E. Hoggan. "The effect of tactile feedback latency in touchscreen interaction." In: *2011 IEEE World Haptics Conference* (2011), pp. 65–70.

[172] Topi Kaaresoja and Stephen Brewster. "Feedback is... late: measuring multimodal delays in mobile device touchscreen interaction." In: *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*. ACM. 2010, p. 2.

[173] Topi Kaaresoja, Stephen Brewster, and Vuokko Lantz. "Towards the Temporally Perfect Virtual Button : Touch-Feedback Simultaneity and Perceived Quality in Mobile Touchscreen Press Interactions." In: *ACM Transactions on Applied Perception* 11.2 (2014).

[174] David Katz. *The world of touch*. Psychology press, 2013.

[175] M Keane and G Dodd. "Subjective assessment of upright piano key vibrations." In: *Acta Acustica united with Acustica* 97.4 (2011), pp. 708–713.

[176] Michael Kennedy and Joyce Bourne. *The concise Oxford dictionary of music*. OUP Oxford, 2004.

[177] Thorsten A Kern. *Engineering haptic devices: a beginner's guide for engineers*. Springer Publishing Company, Incorporated, 2009.

[178] Mitchell L Kietzman and Samuel Sutton. "The interpretation of two-pulse measures of temporal resolution in vision." In: *Vision Research* 8.3 (1968), pp. 287–302.

[179] Lorenz Kilchenmann and Olivier Senn. ""Play in time, but don't play time": Analyzing timing profiles in drum performances." In: *Proceedings of the International Symposium on Performance Science*. 2011, pp. 593–598.

[180] Roberta L Klatzky and Joann Peck. "Please touch: Object properties that invite touch." In: *IEEE Transactions on Haptics* 5.2 (2012), pp. 139–147.

[181] Scott R Klemmer, Björn Hartmann, and Leila Takayama. "How bodies matter: five themes for interaction design." In: *Proceedings of the 6th conference on Designing Interactive systems*. ACM. 2006, pp. 140–149.

[182] Jürgen Konczak, Heidi vander Velden, and Lukas Jaeger. "Learning to play the violin: motor control by freezing, not freeing degrees of freedom." In: *Journal of motor behavior* 41.3 (2009), pp. 243–252.

[183] Katri Kosonen and Roope Raisamo. "Rhythm perception through different modalities." In: *Proc. EuroHaptics*. 2006, pp. 365–370.

[184] Vanessa Krause, Bettina Pollok, and Alfons Schnitzler. "Perception in action: the impact of sensory information on sensorimotor synchronization in musicians and non-musicians." In: *Acta psychologica* 133.1 (2010), pp. 28–37.

[185] Lester E Krueger. "Tactual perception in historical perspective: David Katz's world of touch." In: *Tactual Perception; A Sourcebook* (1982), pp. 1–55.

[186] Gordon Kurtenbach. "Gestures in human-computer communication." In: *The Art of Human-Computer Interface Design, B.* (1990).

[187] Alexandra Kuznetsova, Per B. Brockhoff, and Rune H. B. Christensen. "lmerTest Package: Tests in Linear Mixed Effects Models." In: *Journal of Statistical Software* 82.13 (2017), pp. 1–26.

[188] Tellef Kvifte and Sverre Jensen. *Instruments and the Electronic Age*. Lulu. com, 2007.

[189] Tellef Kvifte and Alexander Refsum Jensenius. "Towards a coherent terminology and model of instrument description and design." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. IRCAM—Centre Pompidou. 2006, pp. 220–225.

[190] Np Lago and Fabio Kon. "The quest for low latency." In: *Proc. of the International Computer Music Conference*. 2004, pp. 33–36.

[191] Jeppe Veirum Larsen and Hendrik Knoche. "Hear You Later Alligator: How delayed auditory feedback affects non-musically trained people's strumming." In: *New Interfaces for Musical Expression*. 2017.

[192] Astrid Twenebowa Larssen, Toni Robertson, and Jenny Edwards. "The feel dimension of technology interaction: exploring tangibles through movement and touch." In: *Proceedings of the 1st international conference on Tangible and embedded interaction*. ACM. 2007, pp. 271–278.

[193] Elisabeth Le Guin. *Boccherini's Body: an essay in carnal musicology*. Univ of California Press, 2005.

[194] Anatole Lécuyer. "Simulating haptic feedback using vision: A survey of research and applications of pseudo-haptic feedback." In: *Presence: Teleoperators and Virtual Environments* 18.1 (2009), pp. 39–53.

[195] Susan J Lederman and Roberta L Klatzky. "Hand movements: A window into haptic object recognition." In: *Cognitive psychology* 19.3 (1987), pp. 342–368.

[196] Marc Leman. *Embodied music cognition and mediation technology*. MIT Press, 2008.

[197] Marc Leman, Frank Desmet, Frederik Styns, Leon Van Noorden, and Dirk Moelants. "Sharing musical expression through embodied listening: a case study based on Chinese guqin music." In: *Music Perception: An Interdisciplinary Journal* 26.3 (2009), pp. 263–278.

[198] James Leonard and Claude Cadoz. "Physical modelling concepts for a collection of multisensory virtual musical instruments." In: *New Interfaces for Musical Expression*. 2015, pp. 150–155.

[199] Giacomo Lepri and Andrew McPherson. "Mirroring the past, from typewriting to interactive art: an approach to the re-design of a vintage technology." In: (2018).

[200] Michael Lester and Jon Boley. "The Effects of Latency on Live Sound Monitoring." In: *Audio Engineering Society Convention 123* (2007), pp. 1–20.

[201] Daniel J Levitin, Stephen McAdams, and Robert L Adams. "Control parameters for musical instruments: a foundation for new mappings of gesture to sound." In: *Organised Sound* 7.2 (2002), pp. 171–189.

[202] Daniel J Levitin, Karon MacLean, Max Mathews, Lonny Chu, Eric Jensen, and Daniel M Dubois. "The perception of cross-modal simultaneity (or "the Greenwich Observatory Problem" revisited)." In: *AIP Conference Proceedings*. Vol. 517. 1. AIP. 2000, pp. 323–329.

[203] Vanessa K Lim, John L Bradshaw, Michael ER Nicholls, and Eckart Altenmüller. "Perceptual differences in sequential stimuli across patients with musician's and writer's cramp." In: *Movement disorders* 18.11 (2003), pp. 1286–1293.

[204] Lian Loke and A Baki Kocaballi. "Choreographic inscriptions: a framework for exploring sociomaterial influences on qualities of movement for HCI." In: *Human Technology* 12 (2016).

[205] Jack M Loomis and Susan J Lederman. "Tactual perception." In: *Handbook of perception and human performances* 2 (1986), p. 2.

[206] Annie Luciani, J-L Florens, and Nicolas Castagné. "From action to sound: a challenging perspective for haptics." In: *Eurohaptics Conference, 2005 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. World Haptics 2005. First Joint*. IEEE. 2005, pp. 592–595.

[207] Annie Luciani, Jean-Loup Florens, Damien Couroussé, and Julien Castet. "Ergotic sounds: A new way to improve playability, believability and presence of virtual musical instruments." In: *Journal of New Music Research* 38.3 (2009), pp. 309–323.

[208] Kevin Lynch. *The image of the city*. Vol. 11. MIT press, 1960.

[209] Donald G MacKay. *The organization of perception and action: A theory for language and other cognitive skills*. Springer Science & Business Media, 2012.

[210] I. Scott MacKenzie and Colin Ware. "Lag as a determinant of human performance in interactive systems." In: *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*. ACM. 1993, pp. 488–493.

[211] Karon E MacLean. "Haptic interaction design for everyday interfaces." In: *Reviews of Human Factors and Ergonomics* 4.1 (2008), pp. 149–194.

[212] Wendy E Mackay. "The interactive thread: exploring methods for multi-disciplinary design." In: *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*. ACM. 2004, pp. 103–112.

[213] Pieter-Jan Maes, Marc Leman, Caroline Palmer, and Marcelo M Wanderley. "Action-based effects on music perception." In: *Frontiers in psychology* 4 (2013).

[214] Thor Magnusson. "Of epistemic tools: Musical instruments as cognitive extensions." In: *Organised Sound* 14.2 (2009), pp. 168–176.

[215] Thor Magnusson. "An epistemic dimension space for musical devices." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. University of Technology Sydney. 2010, pp. 43–46.

[216] Thor Magnusson. "Designing constraints: Composing and performing with digital musical systems." In: *Computer Music Journal* 34.4 (2010), pp. 62–73.

[217] Thor Magnusson. "Musical organics: a heterarchical approach to digital organology." In: *Journal of New Music Research* 46.3 (2017), pp. 286–303.

[218] Thor Magnusson and Enrike Hurtado Mendieta. "The acoustic, the digital and the body: A survey on musical instruments." In: *Proc. of the Conference on New Interfaces for Musical Expression*. ACM. 2007, pp. 94–99.

[219] Teemu Mäki-Patola and Perttu Hämäläinen. "Latency tolerance for gesture controlled continuous sound instrument without tactile feedback." In: *Proc. International Computer Music Conference (ICMC)*. 2004, pp. 1–5.

[220] James C Makous, Robert M Friedman, and Charles J Vierck. "A critical band filter in touch." In: *Journal of Neuroscience* 15.4 (1995), pp. 2808–2818.

[221] Joseph Malloch, David Birnbaum, Elliot Sinyor, and Marcelo M Wanderley. "Towards a new conceptual framework for digital musical instruments." In: *Proceedings of the 9th International Conference on Digital Audio Effects*. 2006, pp. 49–52.

[222] Fiona C Manning and Michael Schutz. "Trained to keep a beat: movement-related enhancements to timing perception in percussionists and non-percussionists." In: *Psychological research* 80.4 (2016), pp. 532–542.

[223] Adnan Marquez-Borbon, Michael Gurevich, A Cavan Fyans, and Paul Stapleton. "Designing digital musical interactions in experimental contexts." In: *contexts* 16.20 (2011), p. 21.

[224] David Marr and A Vision. "A computational investigation into the human representation and processing of visual information." In: *WH San Francisco: Freeman and Company* 1.2 (1982).

[225] Mark T Marshall. "Physical interface design for digital musical instruments." PhD thesis. McGill University, 2009.

[226] Mark T Marshall and Marcelo M Wanderley. "Evaluation of sensors as input devices for computer music interfaces." In: *International Symposium on Computer Music Modeling and Retrieval*. Springer. 2005, pp. 130–139.

[227] Mark T Marshall and Marcelo M Wanderley. "Vibrotactile feedback in digital musical instruments." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. IRCAM—Centre Pompidou. 2006, pp. 226–229.

[228] Mark T Marshall and Marcelo M Wanderley. "Examining the Effects of Embedded Vibrotactile Feedback on the Feel of a Digital Musical Instrument." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2011, pp. 399–404.

[229] Mark T Marshall, Max Hartshorn, Marcelo M Wanderley, and Daniel J Levitin. "Sensor choice for parameter modulations in digital musical instruments: Empirical evidence from pitch modulation." In: *Journal of New Music Research* 38.3 (2009), pp. 241–253.

[230] Dominic W Massaro. "An information-processing analysis of perception and action." In: *Relationships between perception and action*. Springer, 1990, pp. 133–166.

[231] Saúl Maté-Cid. "Vibrotactile perception of musical pitch." PhD thesis. University of Liverpool, 2013.

[232] Francis McGlone and David Reilly. "The cutaneous sensory system." In: *Neuroscience & Biobehavioral Reviews* 34.2 (2010), pp. 148–159.

[233] Harry McGurk and John MacDonald. "Hearing lips and seeing voices." In: *Nature* 264.5588 (1976), pp. 746–748.

[234] Andrew P McPherson and Youngmoo E Kim. "Multidimensional gesture sensing at the piano keyboard." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2011, pp. 2789–2798.

[235] Andrew P McPherson and Youngmoo Kim. "Augmenting the Acoustic Piano with Electromagnetic String Actuation and Continuous Key Position Sensing." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2010, pp. 217–222.

[236] Andrew P McPherson and Victor Zappi. "Exposing the Scaffolding of Digital Instruments with Hardware-Software Feedback Loops." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2015.

[237] Andrew P McPherson, Alan Chamberlain, Adrian Hazzard, Sean McGrath, and Steve Benford. "Designing for exploratory play with a hackable digital musical instrument." In: *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM. 2016, pp. 1233–1245.

[238] Andrew McPherson. "TouchKeys: Capacitive Multi-Touch Sensing on a Physical Keyboard." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2012.

[239] Andrew McPherson, Robert H Jack, and Giulio Moro. "Action-Sound Latency: Are Our Tools Fast Enough?" In: *Proceedings of the international conference on new interfaces for musical expression*. 2016, pp–20.

[240] Andrew McPherson and Victor Zappi. "An environment for submillisecond-latency audio and sensor processing on BeagleBone Black." In: *Audio Engineering Society Convention 138*. Audio Engineering Society. 2015.

[241] Gary E McPherson and John McCormick. "Motivational and self-regulated learning components of musical practice." In: *Bulletin of the Council for Research in Music Education* (1999), pp. 98–102.

[242] Carolina Brum Medeiros and Marcelo M Wanderley. "A comprehensive review of sensors and instrumentation methods in devices for musical expression." In: *Sensors* 14.8 (2014), pp. 13556–13591.

[243] Michael Meehan, Sharif Razzaque, Mary C Whitton, and Frederick P Brooks. "Effect of latency on presence in stressful virtual environments." In: *Virtual Reality, 2003. Proceedings. IEEE*. IEEE. 2003, pp. 141–148.

[244] Sebastian Merchel, Anna Leppin, and Ercan Altinsoy. "Hearing with your body: the influence of whole-body vibrations on loudness perception." In: *Proceedings of the Sixteenth Interna-*

*tional Congress on Sound and Vibration (ICSV16), Kraków, Poland*. 2009.

[245] M Alex Meredith, James W Nemitz, and Barry E Stein. "Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors." In: *Journal of Neuroscience* 7.10 (1987), pp. 3215–3229.

[246] Maurice Merleau-Ponty. *Phenomenology of perception*. Routledge, 2013.

[247] David Merrill, Hayes Raffle, and Roberto Aimi. "The sound of touch: physical manipulation of digital sound." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2008, pp. 739–742.

[248] Romain Michon, Julius O Smith, Matthew Wright, and Chris Chafe. "Augmenting the iPad: the BladeAxe." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2016.

[249] Robert B Miller. "Response time in man-computer conversational transactions." In: *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*. ACM. 1968, pp. 267–277.

[250] Eduardo Reck Miranda and Marcelo M Wanderley. *New digital musical instruments: control and interaction beyond the keyboard*. Vol. 21. AR Editions, Inc., 2006.

[251] Ali Momeni. "Caress: An Enactive Electro-acoustic Percussive Instrument for Caressing Sound." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2015.

[252] F Richard Moore. "The dysfunctions of MIDI." In: *Computer music journal* 12.1 (1988), pp. 19–28.

[253] Fabio Morreale, Andrew P McPherson, and Marcelo M Wanderley. "NIME Identity from the Performer's Perspective." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2018.

[254] Fabio Morreale et al. "Design for Longevity: Ongoing Use of Instruments from NIME 2010-14." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2017.

[255] William Moss and Bryan Cunitz. "Haptic Theremin: Developing a Haptic Musical controller using the Sensable Phantom Omni." In: *International Computer Music Conference*. 2005, 275–277.

[256] Daniel Müllensiefen, Bruno Gingras, Jason Musil, and Lauren Stewart. "The musicality of non-musicians: an index for assessing musical sophistication in the general population." In: *PloS one* 9.2 (2014), e89642.

[257] Saskia K Nagel, Christine Carl, Tobias Kringe, Robert Märtin, and Peter König. "Beyond sensory substitution—learning the sixth sense." In: *Journal of neural engineering* 2.4 (2005), R13.

[258] Chris Nash and Alan Blackwell. "Tracking virtuosity and flow in computer music." In: *Proc. International Computer Music Conference (ICMC)*. 2011.

[259] Chris Nash and Alan Blackwell. "Liveness and Flow in Notation Use." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2012.

[260] Jordi Navarra, Salvador Soto-Faraco, and Charles Spence. "Adaptation to audiotactile asynchrony." In: *Neuroscience letters* 413.1 (2007), pp. 72–76.

[261] Albert Ng, Julian Lepinski, Daniel Wigdor, Steven Sanders, and Paul Dietz. "Designing for low-latency direct-touch input." In: *Proceedings of the 25th annual ACM symposium on User interface software and technology*. ACM. 2012, pp. 453–464.

[262] Albert Ng, Michelle Annett, Paul Dietz, Anoop Gupta, and Walter F Bischof. "In the blink of an eye: investigating latency perception during stylus interaction." In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM. 2014, pp. 1103–1112.

[263] Charles Nichols. "The vBow: development of a virtual violin bow haptic human-computer interface." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. National University of Singapore. 2002, pp. 1–4.

[264] Luc Nijs, Micheline Lesaffre, and Marc Leman. "The musical instrument as a natural extension of the musician." In: *the 5th Conference of Interdisciplinary Musicology*. LAM-Institut jean Le Rond d'Alembert. 2009, pp. 132–133.

[265] Alva Noë. *Action in perception*. MIT press, 2004.

[266] Alva Noë. "What Would Disembodied Music Even Be?" In: *Bodily expression in electronic music: perspectives on reclaiming performativity*. Ed. by Deniz Peters, Gerhard Eckel, and Andreas Dorschel. Routledge, 2012.

[267] Don Norman. *The design of everyday things: Revised and expanded edition*. Basic Books (AZ), 2013.

[268] Donald A Norman. *The invisible computer: why good products can fail, the personal computer is so complex, and information appliances are the solution*. MIT press, 1998.

[269] Donald A Norman. "Natural user interfaces are not natural." In: *interactions* 17.3 (2010), pp. 6–10.

[270] SJ Norman, M Waisvisz, and R Joel. *Touchstone. Catalogue to the first STEIM touch-exhibition*. 1998.

[271] Giacomo Novembre and Peter E Keller. "A conceptual review on action-perception coupling in the musicians' brain: what is it good for?" In: *Frontiers in human neuroscience* 8 (2014), p. 603.

[272] Maura Sile O'Modhrain. "Playing by feel: incorporating haptic feedback into computer-based musical instruments." PhD thesis. Stanford University, 2001.

[273] Sile O'Modhrain. "A framework for the evaluation of digital musical instruments." In: *Computer Music Journal* 35.1 (2011), pp. 28–42.

[274] Sile O'Modhrain and Georg Essl. "PebbleBox and Crumble-Bag: tactile interfaces for granular synthesis." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. National University of Singapore. 2004, pp. 74–79.

[275] Valeria Occelli, Charles Spence, and Massimiliano Zampini. "Audiotactile interactions in temporal perception." In: *Psychonomic bulletin & review* 18.3 (2011), pp. 429–454.

[276] Trond Are Øritsland and Jacob Buur. "Taking the best from a company history-designing with interaction styles." In: *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*. ACM. 2000, pp. 27–38.

[277] Otto Ortmann. *The physiological mechanics of piano technique*. London: Paul Kegan, 1929.

[278] Dan Overholt. "The overtone fiddle: an actuated acoustic instrument." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2011, pp. 4–7.

[279] Sile O'Modhrain and R Brent Gillespie. "Once more, with feeling: Revisiting the role of touch in performer-instrument interaction." In: *Musical Haptics*. Springer, Cham, 2018, pp. 11–27.

[280] Garth Paine. "Towards a Taxonomy of Realtime Interfaces for Electronic Music Performance." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2010, pp. 436–439.

[281] Caroline Palmer, Erik Koopmans, Janeen D Loehr, and Christine Carter. "Movement-related feedback and temporal accuracy in clarinet performance." In: *Music Perception: An Interdisciplinary Journal* 26.5 (2009), pp. 439–449.

[282] Stefano Papetti, Sébastien Schiesser, and Martin Fröhlich. "Multipoint vibrotactile feedback for an expressive musical interface." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2015, pp. 235–240.

[283] Joseph A Paradiso. "The Modular Explosion - Deja Vu or Something New?" In: *Presented at the Voltage Connect Conference* (Mar. 10–11, 2017). Berklee College of Music. Boston MA, 2017.

[284] Laurel S Pardue. "Violin Augmentation Techniques for Learning Assistance." PhD thesis. Queen Mary University of London, 2017.

[285] Richard Parncutt. "Accents and expression in piano performance." In: *Perspektiven und Methoden einer Systemischen Musikwissenschaft* (2003), pp. 163–185.

[286] Mark Paterson. *The senses of touch: Haptics, affects and technologies*. Berg, 2007.

[287] Joann Peck and Terry L Childers. "If I touch it I have to have it: Individual and environmental influences on impulse purchasing." In: *Journal of business research* 59.6 (2006), pp. 765–769.

[288] Joann Peck and Suzanne B Shu. "The effect of mere touch on perceived ownership." In: *Journal of consumer Research* 36.3 (2009), pp. 434–447.

[289] Deniz Peters. "Touch: Apparent, Real, and Absent: On Bodily Expression in Electronic Music." In: *Bodily expression in electronic music: perspectives on reclaiming performativity*. Ed. by Deniz Peters, Gerhard Eckel, and Andreas Dorschel. Routledge, 2012.

[290] P Pfordresher and Caroline Palmer. "Effects of delayed auditory feedback on timing of music performance." In: *Psychological research* 66.1 (2002), pp. 71–79.

[291] Peter Q Pfordresher. "Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing." In: *Journal of Experimental Psychology: Human Perception and Performance* 29.5 (2003), p. 949.

[292] Peter Q Pfordresher. "Auditory feedback in music performance: the role of melodic structure and musical skill." In: *Journal of Experimental Psychology: Human Perception and Performance* 31.6 (2005), p. 1331.

[293] Peter Q Pfordresher. "Coordination of perception and action in music performance." In: *Advances in Cognitive Psychology* 2.2-3 (2006), pp. 183–198.

[294] Peter Q Pfordresher. "Auditory feedback in music performance: The role of transition-based similarity." In: *Journal of Experimental Psychology: Human Perception and Performance* 34.3 (2008), p. 708.

[295] Peter Q Pfordresher and Simone Dalla Bella. "Delayed auditory feedback and movement." In: *Journal of Experimental Psychology: Human Perception and Performance* 37.2 (2011), p. 566.

[296] Peter Q Pfordresher and Caroline Palmer. "Effects of hearing the past, present, or future during music performance." In: *Attention, Perception, & Psychophysics* 68.3 (2006), pp. 362–376.

[297] Martin Pielot, Anastasia Kazakova, Tobias Hesselmann, Wilko Heuten, and Susanne Boll. "PocketMenu: non-visual menus for touch screen devices." In: *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services*. ACM. 2012, pp. 327–330.

[298] Jon Pigrem and Andrew McPherson. "Do We Speak Sensor? Cultural Constraints of Embodied Interaction." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2018.

[299] Betina Piqueras-Fiszman and Charles Spence. "The weight of the bottle as a possible extrinsic cue with which to estimate the price (and quality) of the wine? Observed correlations." In: *Food Quality and Preference* 25.1 (2012), pp. 41–45.

[300] Miller S Puckette et al. "Pure Data." In: *Proc. International Computer Music Conference (ICMC)*. 1997.

[301] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2017. URL: https://www.R-project.org/.

[302] Thomas Rammsayer and Eckart Altenmüller. "Temporal information processing in musicians and nonmusicians." In: *Music Perception: An Interdisciplinary Journal* 24.1 (2006), pp. 37–48.

[303] Matthew Ratcliffe. "Touch and the sense of reality." In: *The hand: An organ of the mind*. Ed. by Zdravko Radman. 1st ed. MIT Press, 2013.

[304] Pedro Rebelo. "Haptic sensation and instrumental transgression." In: *Contemporary Music Review* 25.1-2 (2006), pp. 27–35.

[305] Bruno H Repp. "Effects of auditory feedback deprivation on expressive piano performance." In: *Music Perception: An Interdisciplinary Journal* 16.4 (1999), pp. 409–438.

[306] Bruno H Repp. "Compensation for subliminal timing perturbations in perceptual-motor synchronization." In: *Psychological research* 63.2 (2000), pp. 106–128.

[307] Bruno H Repp and Rebecca Doggett. *Tapping to a Very Slow Beat: A Comparison of Musicians and Nonmusicians*. 2007.

[308] Bruno H Repp and Yi Huang Su. "Sensorimotor synchronization: a review of recent research (2006–2012)." In: *Psychonomic bulletin & review* 20.3 (2013), pp. 403–452.

[309] Paul Bach-y Rita and Stephen W Kercel. "Sensory substitution and the human–machine interface." In: *Trends in cognitive sciences* 7.12 (2003), pp. 541–546.

[310] Giacomo Rizzolatti and Laila Craighero. "The mirror-neuron system." In: *Annu. Rev. Neurosci.* 27 (2004), pp. 169–192.

[311] Davide Rocchesso, Stefano Delle Monache, and Stefano Papetti. "Multisensory texture exploration at the tip of the pen." In: *International Journal of Human-Computer Studies* 85 (2016), pp. 47–56.

[312] Davide Rocchesso and Federico Fontana. *The sounding object*. Mondo estremo, 2003.

[313] Pablo Rojas. "To become one with the instrument: The unfolding of a musical topography." In: *Culture & Psychology* 21.2 (2015), pp. 207–230.

[314] David A Rosenbaum. *Human motor control*. Academic press, 2009.

[315] Joseph Rovan and Vincent Hayward. "Typology of tactile sounds and their synthesis in gesture-driven computer music performance." In: *Trends in gestural control of music* (2000), pp. 297–320.

[316] Frank A Russo, Paolo Ammirante, and Deborah I Fels. "Vibrotactile discrimination of musical timbre." In: *Journal of Experimental Psychology: Human Perception and Performance* 38.4 (2012), p. 822.

[317] Charalampos Saitis, Bruno L Giordano, Claudia Fritz, and Gary P Scavone. "Perceptual evaluation of violins: A quantitative analysis of preference judgments by experienced players." In: *The Journal of the Acoustical Society of America* 132.6 (2012), pp. 4002–4012.

[318] Charalampos Saitis, Claudia Fritz, Gary P Scavone, Catherine Guastavino, and Danièle Dubois. "Perceptual evaluation of violins: A psycholinguistic analysis of preference verbal descriptions by experienced musicians." In: *The Journal of the Acoustical Society of America* 141.4 (2017), pp. 2746–2757.

[319] Christine Satchell and Paul Dourish. "Beyond the user: use and non-use in HCI." In: *Proceedings of the 21st annual conference of the Australian computer-human interaction special interest group: Design: Open 24/7*. ACM. 2009, pp. 9–16.

[320] Michael Schaefer, Hans-Jochen Heinze, and Michael Rotte. "Embodied empathy for tactile events: interindividual differences and vicarious somatosensory responses during touch observation." In: *Neuroimage* 60.2 (2012), pp. 952–957.

[321] Pierre Schaeffer. *In search of a concrete music*. Vol. 15. Univ of California Press, 2012.

[322] Bert Schiettecatte and Jean Vanderdonckt. "AudioCubes: a distributed cube tangible interface based on interaction range for sound design." In: *Proceedings of the 2nd international conference on Tangible and embedded interaction*. ACM. 2008, pp. 3–10.

[323] Daniel Schlessinger and Julius O Smith. "The Kalichord: A Physically Modeled Electro-Acoustic Plucked String Instrument." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2009, pp. 98–101.

[324] W Andrew Schloss. "Using contemporary technology in live performance: The dilemma of the performer." In: *Journal of New Music Research* 32.3 (2003), pp. 239–242.

[325] Benjamin G Schultz and Floris T van Vugt. "Tap Arduino: An Arduino microcontroller for low-latency auditory feedback in sensorimotor synchronization experiments." In: *Behavior research methods* 48.4 (2016), pp. 1591–1607.

[326] Marlon Schumacher, Marcello Giordano, Marcelo M Wanderley, and Sean Ferguson. "Vibrotactile notification for live electronics performance: A prototype system." In: *Proceedings of the International Symposium on Computer Music Multidisciplinary Research (CMMR)*. 2013, pp. 516–525.

[327] Martin Schürmann, Gina Caetano, Yevhen Hlushchuk, Veikko Jousmäki, and Riitta Hari. "Touch activates human auditory cortex." In: *Neuroimage* 30.4 (2006), pp. 1325–1331.

[328] Orit Shaer and Eva Hornecker. "Tangible user interfaces: past, present, and future directions." In: *Foundations and Trends in Human-Computer Interaction* 3.1–2 (2010), pp. 1–137.

[329] Orit Shaer and Robert JK Jacob. "A specification paradigm for the design and implementation of tangible user interfaces." In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 16.4 (2009), p. 20.

[330] Lawrence Shapiro. *Embodied cognition*. Routledge, 2010.

[331] Eric Sheffield and Michael Gurevich. "Distributed Mechanical Actuation of Percussion Instruments." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. The School of Music, the Center for Computation, and Technology (CCT), Louisiana State University. 2015, pp. 11–15.

[332] Patrick Shove and Bruno H Repp. "Musical motion and performance: theoretical and empirical perspectives." In: *The practice of performance: Studies in musical interpretation* (1995), pp. 55–83.

[333] John A Sloboda, Jane W Davidson, Michael JA Howe, and Derek G Moore. "The role of practice in the development of performing musicians." In: *British journal of psychology* 87.2 (1996), pp. 287–309.

[334] Christopher Small. *Musicking: The meanings of performing and listening*. Wesleyan University Press, 2011.

[335] Mary M Smyth and Peter E Morris. *Cognition in action*. Psychology Press, 1994.

[336] Salvador Soto-Faraco and Gustavo Deco. "Multisensory contributions to the perception of vibrotactile events." In: *Behavioural brain research* 196.2 (2009), pp. 145–154.

[337] Charles Spence. "Crossmodal correspondences: A tutorial review." In: *Attention, Perception, & Psychophysics* 73.4 (2011), pp. 971–995.

[338] Charles Spence and Alberto Gallace. "Multisensory design: Reaching out to touch the consumer." In: *Psychology & Marketing* 28.3 (2011), pp. 267–308.

[339] Charles Spence and Cesare Parise. "Prior-entry: A review." In: *Consciousness and cognition* 19.1 (2010), pp. 364–379.

[340] MC Marieke van der Steen, EBD Molendijk, E Altenmüller, and S Furuya. "Expert pianists do not listen: the expertise-dependent influence of temporal perturbation on the production of sequential movements." In: *Neuroscience* 269 (2014), pp. 290–298.

[341] C Steenson, M Rodger, and W Matthew. "Bringing sounds into use: thinking of sounds as materials and a sketch of auditory affordances." In: *Open Psychol. J* 8 (2015), pp. 174–182.

[342] Barry E Stein. *The new handbook of multisensory processing*. MIT Press, 2012.

[343] Prisca Stenneken, Wolfgang Prinz, Jonathan Cole, Jacques Paillard, and Gisa Aschersleben. "The effect of sensory feedback on the timing of movements: evidence from deafferented patients." In: *Brain research* 1084.1 (2006), pp. 123–131.

[344] Dan Stowell, Mark D Plumbley, and Nick Bryan-Kinns. "Discourse Analysis Evaluation Method for Expressive Musical Interfaces." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2008, pp. 81–86.

[345] Philip Strain, A Dawn Shaikh, and Richard Boardman. "Thinking but not seeing: think-aloud for non-sighted users." In: *CHI'07 Extended Abstracts on Human Factors in Computing Systems*. ACM. 2007, pp. 1851–1856.

[346] Paul Strohmeier, Jesse Burstyn, Juan Pablo Carrascal, Vincent Levesque, and Roel Vertegaal. "ReFlex: A Flexible Smartphone with Active Haptic Feedback for Bend Input." In: *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM. 2016, pp. 185–192.

[347] Bob Sturm, Joao Felipe Santos, and Iryna Korshunova. "Folk music style modelling by recurrent neural networks with long short term memory units." In: *Proc. ISMIR*. 2015.

[348] David Sudnow. *Ways of the hand: The organization of improvised conduct*. MIT Press, 1978.

[349] Dag Svanæs. “Interaction design for and with the lived body: Some implications of merleau-ponty’s phenomenology.” In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 20.1 (2013), p. 8.

[350] Ana Tajadura-Jiménez, Bin Liu, Nadia Bianchi-Berthouze, and Frédéric Bevilacqua. “Using sound in multi-touch interfaces to change materiality and touch behavior.” In: *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. ACM. 2014, pp. 199–202.

[351] Marc R Thompson and Geoff Luck. “Exploring relationships between pianists’ body movements, their expressive intentions, and structural elements of the music.” In: *Musicae Scientiae* 16.1 (2012), pp. 19–40.

[352] Adam R Tindale, Ajay Kapur, George Tzanetakis, Peter Driessen, and Andrew Schloss. “A comparison of sensor strategies for capturing percussive gestures.” In: *Proceedings of the Conference on New Interfaces for Musical Expression*. National University of Singapore. 2005, pp. 200–203.

[353] John Tresch and Emily I Dolan. “Toward a new organology: Instruments of music and science.” In: *Osiris* 28.1 (2013), pp. 278–298.

[354] Chia-Jung Tsay. “Sight over sound in the judgment of music performance.” In: *Proceedings of the National Academy of Sciences* 110.36 (2013), pp. 14580–14585.

[355] Kai Tuuri, Jaana Parviainen, and Antti Pirhonen. “Who controls who? Embodied control within human–technology choreographies.” In: *Interacting with Computers* 29.4 (2017), pp. 494–511.

[356] Kai Tuuri, Antti Pirhonen, and Eve Hoggan. “Some severe deficiencies of the input-output HCI-paradigm and their influence on practical design.” In: *Designing Beyond the Product* (2009).

[357] George Tzanetakis, Sidney Fels, and Michael Lyons. "Blending the physical and the virtual in music technology: from interface design to multi-modal signal processing." In: *Proceedings of the 21st ACM international conference on Multimedia*. ACM. 2013, pp. 1119–1120.

[358] Brygg Ullmer and Hiroshi Ishii. "Emerging frameworks for tangible user interfaces." In: *IBM systems journal* 39.3.4 (2000), pp. 915–931.

[359] Maaike Van Den Haak, Menno De Jong, and Peter Jan Schellens. "Retrospective vs. concurrent think-aloud protocols: testing the usability of an online library catalogue." In: *Behaviour & information technology* 22.5 (2003), pp. 339–351.

[360] Doug Van Nort. "Instrumental listening: sonic gesture as design principle." In: *Organised sound* 14.2 (2009), pp. 177–187.

[361] Francisco J Varela, Eleanor Rosch, and Evan Thompson. "The Embodied Mind: Cognitive Science and Human Experience." In: (1992).

[362] Bill Verplank. "Interaction design sketchbook." In: *Unpublished paper for CCRMA course Music 250a* (2003).

[363] Bill Verplank, Michael Gurevich, and Max Mathews. "The Plank: designing a simple haptic controller." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. National University of Singapore. 2002, pp. 1–4.

[364] Bill Verplank, Craig Sapp, and Max Mathews. "A course on controllers." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. National University of Singapore. 2001, pp. 1–4.

[365] Roel Vertegaal, Tamas Ungvary, and Michael Kieslinger. "Towards a musician's cockpit: Transducers, feedback and musical function." In: *Proc. International Computer Music Conference (ICMC)*. Vol. 96. 1996, pp. 308–311.

[366] Jean Vroomen and Mirjam Keetels. "Perception of intersensory synchrony: a tutorial review." In: *Attention, Perception, & Psychophysics* 72.4 (2010), pp. 871–884.

[367] Marcelo M Wanderley and Philippe Depalle. "Gestural control of sound synthesis." In: *Proceedings of the IEEE* 92.4 (2004), pp. 632–644.

[368] Marcelo M Wanderley, Bradley W Vines, Neil Middleton, Cory McKay, and Wesley Hatch. "The musical significance of clarinetists' ancillary gestures: An exploration of the field." In: *Journal of New Music Research* 34.1 (2005), pp. 97–113.

[369] Marcelo Mortensen Wanderley and Nicola Orio. "Evaluation of input devices for musical expression: Borrowing tools from hci." In: *Computer Music Journal* 26.3 (2002), pp. 62–76.

[370] Nicholas Ward. "Effortful Interaction: A New Paradigm for the Design of Digital Musical Instruments." PhD thesis. Queen's University Belfast, 2013.

[371] Nicholas Ward, Kedzie Penfield, Sile O'Modhrain, and R Benjamin Knapp. "A Study of Two Thereminists: Towards Movement Informed Instrument Design." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2008, pp. 117–121.

[372] William H Warren. "Perceiving affordances: visual guidance of stair climbing." In: *Journal of experimental psychology: Human perception and performance* 10.5 (1984), p. 683.

[373] David Wessel. "An enactive approach to computer music performance." In: *Le Feedback dans la Creation Musical, Lyon, France* (2006), pp. 93–98.

[374] David Wessel and Matthew Wright. "Problems and Prospects for Intimate Musical Control of Computers." In: *Computer Music Journal* 26.3 (2002), pp. 11–14.

[375] David Wessel, Matthew Wright, and John Schott. "Intimate musical control of computers with a variety of controllers and gesture mapping metaphors." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. National University of Singapore. 2002, pp. 1–3.

[376] Mikael Wiberg. "Methodology for materiality: interaction design research through a material lens." In: *Personal and ubiquitous computing* 18.3 (2014), pp. 625–636.

[377] Norbert Wiener. *Cybernetics or Control and Communication in the Animal and the Machine*. Vol. 25. MIT press, 1961.

[378] John R Wilson. "Fundamentals of ergonomics in theory and practice." In: *Applied ergonomics* 31.6 (2000), pp. 557–567.

[379] Indiana Wollman, Claudia Fritz, and Jacques Poitevineau. "Influence of vibrotactile feedback on some perceptual features of violins." In: *The Journal of the Acoustical Society of America* 136.2 (2014), pp. 910–921.

[380] Indiana Wollman, Claudia Fritz, Jacques Poitevineau, and Stephen McAdams. "Investigating the role of auditory and tactile modalities in violin quality evaluation." In: *PloS one* 9.12 (2014), e112552.

[381] Daniel M Wolpert, Zoubin Ghahramani, and Michael I Jordan. "An internal model for sensorimotor integration." In: *Science* 269.5232 (1995), pp. 1880–1882.

[382] Matthew Wright, Ryan J Cassidy, and Michael Zbyszynski. "Audio and Gesture Latency Measurements on Linux and OSX." In: *Proc. International Computer Music Conference (ICMC)*. 2004.

[383] Matthew Wright, Adrian Freed, et al. "Open SoundControl: A New Protocol for Communicating with Sound Synthesizers." In: *Proc. International Computer Music Conference (ICMC)*. 1997.

[384] Haijun Xia, Ricardo Jota, Benjamin McCanny, Zhe Yu, Clifton Forlines, Karan Singh, and Daniel Wigdor. "Zero-latency tapping: using hover information to predict touch locations and eliminate touchdown latency." In: *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM. 2014, pp. 205–214.

[385] Shiyan Yang, Kathryn Tippey, and Thomas K Ferris. "Exploring the Emergent Perception of Haptic Beats from Paired Vibrotactile Presentation." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 58. 1. SAGE Publications Sage CA: Los Angeles, CA. 2014, pp. 1716–1720.

[386] Aubrey J Yates. "Recent empirical and theoretical approaches to the experimental manipulation of speech in normal subjects and in stammerers." In: *Behaviour research and therapy* 1.2 (1963), pp. 95–119.

[387] Yongjae Yoo and Seungmoon Choi. "An Initial Study on Pitch Correction Guidance for String Instruments Using Haptic Feedback." In: *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer. 2014, pp. 241–247.

[388] Yongjae Yoo and Seungmoon Choi. "A longitudinal study of haptic pitch correction guidance for string instrument players." In: *World Haptics Conference (WHC), 2017 IEEE*. IEEE. 2017, pp. 177–182.

[389] Victor Zappi and Andrew McPherson. "Dimensionality and Appropriation in Digital Musical Instrument Design." In: *Proceedings of the Conference on New Interfaces for Musical Expression*. 2014, pp. 455–460.

[390] Robert J Zatorre, Joyce L Chen, and Virginia B Penhune. "When the brain plays music: auditory–motor interactions in music perception and production." In: *Nature reviews neuroscience* 8.7 (2007), p. 547.

[391] Simone Zimmermann, Sonja Rümelin, and Andreas Butz. "I feel it in my fingers: haptic guidance on touch surfaces." In: *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*. ACM. 2014, pp. 9–12.

[392] Oren Zuckerman, Tamar Gal, Tal Keren-Capelovitch, Tal Karsovsky, Ayelet Gal-Oz, and Patrice L Tamar Weiss. "DataSpoon: Overcoming Design Challenges in Tangible and Embedded Assistive Technologies." In: *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM. 2016, pp. 30–37.

# A

# APPENDIX

This appendix contains the questionnaires used in each study.

QUESTIONNAIRE FROM THE STUDY IN CHAPTER 4 Each of the following questions were answered with a rating from 1 to 10 with 1 representing 'Not at all' and 10 representing 'Very Much'.

- How successfully did you play in tune?
- How hard was it to play in tune?
- Were you able to maintain your desired tempo?
- How mentally demanding was the tactile feedback?
- How much did the tactile feedback assist tuning?
- Which was your preferred condition? (choice of all 4 conditions)

**Age**

[Please choose]

**Gender**

[Please choose]

**Musical Background**

Which of the following would you use to define yourself?

- Instrumentalist
- Singer
- Musician
- Composer
- Sound artist
- Audio engineer
- Electronic musician
- Luthier
- Sound designer
- Producer
- Conductor

**What level of music education have you received?**

- Self taught
- Music lessons to high school level
- Bachelors degree in music
- Masters level degree in music
- Doctorate in music
- Private lessons

  For how many years?
- Other

**What is your main instrument?**

**How many years have you played it for?**

**What other instruments do you play and to what level?**

**Do you use a computer in your musical practice? If so, how?**

[Please choose]

**What musical genre(s) best describe your musical style or preference?**

Back

Next

Figure A.1: The exit survey from the study in Chapter 4 and Chapter 5.

19/03/2019 Questionnaire

Testing the Online Survey Software SoSci Survey

Queen Mary
University of London

## Pair 1

**Responsiveness**
How responsive did you find the instrument under each setting?

A is much more responsive than B | B is much more responsive than A

**Naturalness**
How natural was your interaction with the instrument under each setting?

A is much more natural than B | B is much more natural than A

**Temporal control**
To what extent were you able to control your timing with the settings?

A is much easier to control the timing of than B | B is much easier to control the timing of than A

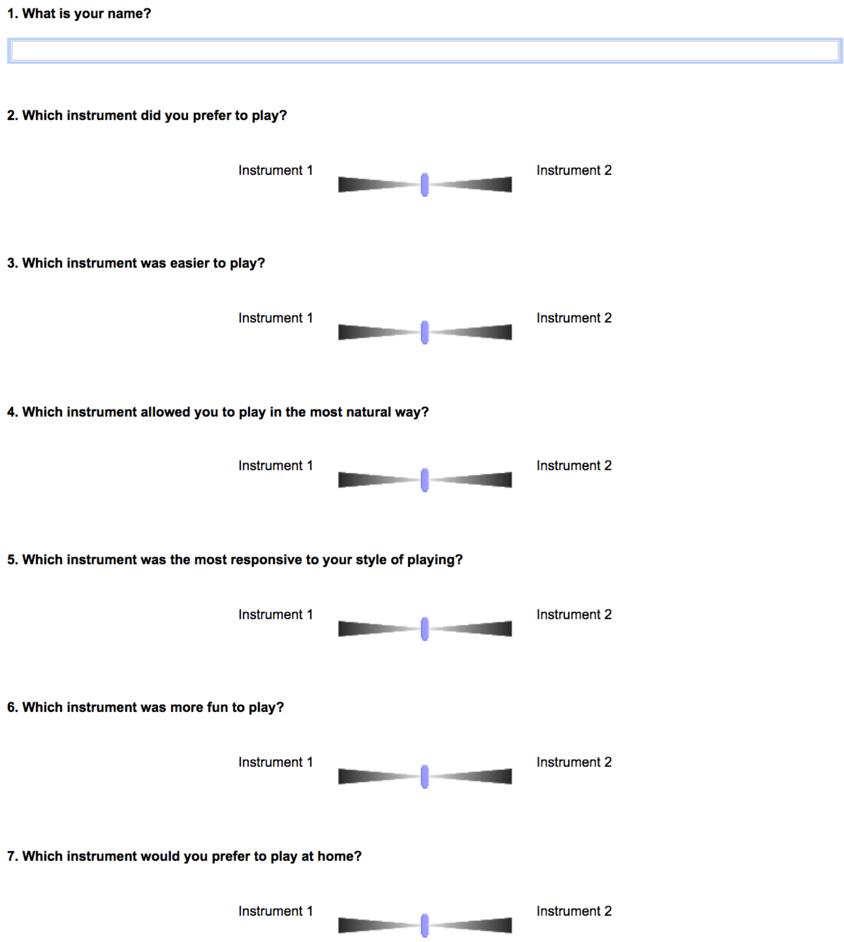
**General preference**
What was your general preference of setting?

A is much more preferable than B | B is much more preferable than A

Back Next

https://www.soscisurvey.de/test161315/index.php?i=RAM3INORNI3E&rnd=ZLIU 1/2

Figure A.2: The questionnaire that was answered for each pair of conditions in the study presented in Chapter 5.



**1. What is your name?**

**2. Which instrument did you prefer to play?**

Instrument 1 Instrument 2

**3. Which instrument was easier to play?**

Instrument 1 Instrument 2

**4. Which instrument allowed you to play in the most natural way?**

Instrument 1 Instrument 2

**5. Which instrument was the most responsive to your style of playing?**

Instrument 1 Instrument 2

**6. Which instrument was more fun to play?**

Instrument 1 Instrument 2

**7. Which instrument would you prefer to play at home?**

Instrument 1 Instrument 2

Figure A.3: The first page of the questionnaire that was answered for each pair of conditions in the study presented in Chapter 6.

**8. Which instrument could you imagine playing in a folk session?**

Instrument 1 Instrument 2

**9. Were the instruments similar to any other instruments you have played?**

Instrument 1

Instrument 2

**10. Which instrument was most similar to a guitar?**

Instrument 1 Instrument 2

**11. How well did you play the accompaniment on each instrument?**

| | Not at all well | | | | Very well |
|---|---|---|---|---|---|
| Instrument 1 | | | | | |
| Instrument 2 | | | | | |

**12. Did the instrument sound like you expected?**

| | Not at all | | | | Exactly as expected |
|---|---|---|---|---|---|
| Instrument 1 | | | | | |
| Instrument 2 | | | | | |

**13. How experienced are you at playing the guitar?**

Figure A.4: The second page of the questionnaire that was answered for each pair of conditions in the study presented in Chapter 6.

## COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/